An introduction to Riemannian geometry

Preliminary version 2

Claudio Gorodski

March 8, 2022

Contents

Preliminaries

0.1 Introduction

The richness of Riemannian geometry is that it has many ramifications and connections to other fields in mathematics and physics. Probably by the very same reasons, it requires quite a lot of language and machinery to get going. In this chapter, we assemble a collection of results and techniques about smooth manifolds and vector fields that we will use in later chapters to develop the theory. Most of the proofs are given and in other cases references are supplied. Despite that, the pace is quick and the absolute beginner is strongly encouraged to supplement the text with other sources.

0.2 Smooth manifolds

The theory of smooth manifolds is a natural and very useful generalization of the differential calculus on $\mathbf{R}^n$. Namely, a smooth manifold is an object that, in the small, looks like a piece of Euclidean space. More formally, a *smooth manifold of dimension n* is a topological space M that can be covered by open sets $\{U_\alpha\}_\alpha$, each of which is homeomorphic to an open subset of Euclidean space under a map $\varphi_\alpha : U_\alpha \to \mathbf{R}^n$; the pair $(U_\alpha, \varphi_\alpha)$ is called a *local chart*; moreover, the following important compatibility condition is required: the *transition maps*

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

must be smooth for all α, β. The family $\{(U_\alpha, \varphi_\alpha)\}_\alpha$ is called a *smooth atlas*. For technical reasons, one also requires that M be Hausdorff and second-countable, and that the smooth atlas $\{(U_\alpha, \varphi_\alpha)\}_\alpha$ be maximal. The basic idea behind this definition is that one can carry some notions and results of differential calculus on $\mathbf{R}^n$ to smooth manifolds via the local charts, the compatibility condition ensuring well defined objects.

A local chart $\varphi : U \to \mathbf{R}^n$ has as components functions usually denoted $x_i : U \to \mathbf{R}$. In this way, a local chart $\varphi = (x_1, \ldots, x_n) : U \to \mathbf{R}^n$ is sometimes also called a *system of local coordinates*, and a transition map is called a *change of local coordinates*.

0.2.1 Examples (First examples of smooth manifolds)

(a) Of course, $\mathbf{R}^n$ is a smooth manifold with the identity map as chart. More generally, any real vector space is a smooth manifold, simply by choosing a basis and identifying with $\mathbf{R}^n$.

(b) An open subset U of a smooth manifold M is also a smooth submanifold: one restricts the local charts of M to U.

(c) The product $M \times N$ of smooth manifolds, with the product topology, is naturally a smooth manifold: typical charts have the form $\varphi_\alpha \times \psi_\beta : U_\alpha \times V_\beta \to \mathbf{R}^m \times \mathbf{R}^n = \mathbf{R}^{m+n}$, where $\varphi_\alpha : U_\alpha \to \mathbf{R}^m$, $\psi_\beta : U_\beta \to \mathbf{R}^n$ are charts of M, N, respectively. Note that $\dim M \times N = \dim M + \dim N$.

(d) It follows from (a) and (b) that the group $GL(n, \mathbf{R})$ of invertible real matrices of size n is a smooth manifold. ⋆

Embedded submanifolds

Let N be a smooth manifold of dimension $n+k$. A subset M of N is called an *embedded submanifold* of N of dimension n if M has the topology induced from N and, for every $p \in M$, there exists a local chart (U, φ) of N with $p \in U$ such that $\varphi(U \cap M) = \varphi(U) \cap \mathbf{R}^n$, where we view $\mathbf{R}^n$ as a subspace of $\mathbf{R}^{n+k}$ in the standard way. We say that (U, φ) is a local chart of N *adapted* to M. Note that in this case the adapted chart (U, φ) induces a local chart $(U \cap M, \varphi|_{U \cap M})$ of M so that M is a smooth manifold in its own right (here the compatibility conditions for the local charts of M follow from those for the local charts of N adapted to M).

0.2.2 Examples (Examples of embedded submanifolds)

(a) An open subset of a smooth manifold is an embedded submanifold of the same dimension.

(b) The graph of a smooth mapping $f : U \to \mathbf{R}^m$, where U is an open subset of $\mathbf{R}^n$, is a smooth submanifold of $\mathbf{R}^{n+m}$ of dimension n. In fact, an adapted local chart is given by $\varphi : U \times \mathbf{R}^m \to U \times \mathbf{R}^m$, $\varphi(p, q) = (p, q - f(p))$, where $p \in \mathbf{R}^m$ and $q \in \mathbf{R}^n$. More generally, if a subset M of $\mathbf{R}^{m+n}$ can be covered by open sets each of which is the graph of a smooth mapping from an open subset of $\mathbf{R}^n$ into $\mathbf{R}^m$, then M is an embedded submanifold of $\mathbf{R}^{n+m}$.

(c) It follows from (b) that the *n-sphere*

$$S^n = \{ (x_1, \ldots, x_{n+1}) \mid x_1^2 + \cdots x_{n+1}^2 = 1 \}$$

is an n-dimensional embedded submanifold of $\mathbf{R}^{n+1}$.

(d) The product of n-copies of the circle S^1 is a n-dimensional manifold called the *n-torus* and denoted by T^n. ⋆

Smooth mappings

A smooth mapping between two smooth manifolds is defined to be a continuous mapping whose local representations with respect to charts on both manifolds is smooth. Namely, let M and N be two smooth manifolds and let $\Omega \subset M$ be open. A continuous map $f : \Omega \to N$ is called *smooth* if and only if

$$\psi \circ f \circ \varphi^{-1} : \varphi(\Omega \cap U) \to \psi(V)$$

is smooth as a map between open sets of Euclidean spaces, for every local charts (U, φ) of M and (V, ψ) of N. Clearly, the composition of two smooth maps is again smooth.

A smooth map $f : M \to N$ between smooth manifolds is called a *diffeomorphism* if it is invertible and the inverse $f^{-1} : N \to M$ is also smooth. Also, $f : M \to N$ is called a *local diffeomorphism* if every $p \in M$ admits an open neighborhood U such that $f(U)$ is open and f defines a diffeomorphism from U onto $f(U)$.

Tangent space and differential

Since arbitrary smooth manifolds in principle do not come with an embedding into an Euclidean space, the tangent space must be constructed abstractly. The philosophy amounts to use the "differential" (not yet defined) of the local charts of M around p to model the tangent space at p.

Let M be a smooth manifold of dimension n, and let $p \in M$. The *tangent space* of M at p is the set T_pM of all pairs (a, φ) — where $a \in \mathbf{R}^n$ and (U, φ) is a local chart around p — quotiented by the equivalence relation

$$(a, \varphi) \sim (b, \psi) \quad \text{if and only if} \quad d(\psi \circ \varphi^{-1})_{\varphi(p)}(a) = b.$$

It follows form the chain rule in $\mathbf{R}^n$ that this is indeed an equivalence relation, and we denote the equivalence class of (a, φ) be $[a, \varphi]$. Each such equivalence class is called a *tangent vector* at p. For a fixed local chart (U, φ) around p, the map

$$a \in \mathbf{R}^n \mapsto [a, \varphi] \in T_pM$$

is a bijection, and it follows from the linearity of $d(\psi \circ \varphi^{-1})_{\varphi(p)}$ that we can use it to transfer the vector space structure of $\mathbf{R}^n$ to T_pM. Note that $\dim T_pM = \dim M$.

Let $(U, \varphi = (x_1, \ldots, x_n))$ be a local chart of M, and denote by $\{e_1, \ldots, e_n\}$ the canonical basis of $\mathbf{R}^n$. The *coordinate vectors* at p are with respect to this chart are defined to be

$$\left.\frac{\partial}{\partial x_i}\right|_p = [e_i, \varphi].$$

Note that

$$\left\{ \left.\frac{\partial}{\partial x_1}\right|_p, \ldots, \left.\frac{\partial}{\partial x_n}\right|_p \right\} \tag{0.2.3}$$

is a basis of T_pM.

In the case of $\mathbf{R}^n$, for each $p \in \mathbf{R}^n$ there is a canonical isomorphism $\mathbf{R}^n \to T_p\mathbf{R}^n$ given by

$$a \mapsto [a, \mathrm{id}], \tag{0.2.4}$$

where id is the identity map of $\mathbf{R}^n$. Usually we will make this identification without further comment. In particular, $T_p\mathbf{R}^n$ and $T_q\mathbf{R}^n$ are canonically isomorphic for every p, $q \in \mathbf{R}^n$. In the case of a general smooth manifold M, obviously there are no such canonical isomorphisms.

Next, let $f : M \to N$ be a smooth map between smooth manifolds. Fix a point $p \in M$, and local charts (U, φ) of M around p, and (V, ψ) of N around $q = f(p)$. The *differential of f at p* is the linear map

$$df_p : T_pM \to T_qN$$

given by

$$[a, \varphi] \mapsto [d(\psi \circ f \circ \varphi^{-1})_{\varphi(p)}(a), \psi].$$

It is easy to check that this definition does not depend on the choices of local charts. Using the identification (0.2.4), one checks that $d\varphi_p : T_pM \to \mathbf{R}^n$ and $d\psi_q : T_pM \to \mathbf{R}^n$ are linear isomorphisms and

$$df_p = (d\psi_q)^{-1} \circ d(\psi \circ f \circ \varphi^{-1})_{\varphi(p)} \circ d\varphi_p.$$

It is also a simple exercise to prove the following important proposition.

0.2.5 Proposition (Chain rule) *Let M, N, P be smooth manifolds. If $f : M \to N$ and $g : N \to P$ are smooth maps, then $g \circ f : M \to P$ is a smooth map and*

$$d(g \circ f)_p = dg_{f(p)} \circ df_p$$

for $p \in M$.

Consider now the case of a smooth map $f : M \to \mathbf{R}$. Then $df_p : T_pM \to T_{f(p)}\mathbf{R} \cong \mathbf{R}$. For $v \in T_pM$, the number

$$v(f) = df_p(v)$$

is called the *directional derivative* of f with respect to v. Fix a coordinate chart $(U, \varphi = (x_1, \ldots, x_n))$ around p and apply this to $f = x_i$. Since $x_j \circ \varphi^{-1} : \varphi(U) \to \mathbf{R}$ is just the restriction of the linear projection onto the jth coordinate of $\mathbf{R}^n$, for any $v = \sum_{i=1}^n a_i \frac{\partial}{\partial x_i}|_p$, we have

$$v(x_j) = d(x_j)_p(v) = d(x_j \circ \varphi^{-1})_{\varphi(p)} \left(\sum_{i=1}^n a_i e_i \right) = a_j,$$

showing that

$$\{dx_1|_p, \ldots, dx_n|_p\}$$

is the basis of T_pM^* dual of the basis (0.2.3).

Finally, a *smooth curve* in M is simply a smooth map $\gamma : (a, b) \to M$ where (a, b) is an interval of $\mathbf{R}$. One can also consider smooth curves γ in M defined on a closed interval $[a, b]$. This simply means that γ admits a smooth extension to an open interval $(a - \epsilon, b + \epsilon)$ for some $\epsilon > 0$. If $\gamma : (a, b) \to M$ is a smooth curve, the *tangent vector* to γ at $t \in (a, b)$ is

$$\dot\gamma(t) = d\gamma_t(e_1) \in T_{\gamma(t)}M,$$

where $e_1 = 1 \in \mathbf{R}$.

Tangent and cotangent bundles

There is a situation in which we want to endow a set X with no natural topology with a structure of smooth manifold. In that case there is a way of using charts to define the topology and smooth structure simultaneously. Namely, fix an integer n, and let $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ be a countable covering of X by arbitrary subsets, on each of which is defined a bijective map $\varphi_\alpha : U_\alpha \to \mathbf{R}^n$ onto an open subset of $\mathbf{R}^n$ such that the sets $\varphi_\alpha(U_\alpha \cap U_\beta)$, $\varphi_\beta(U_\alpha \cap U_\beta)$ are open in $\mathbf{R}^n$ and the transition maps $\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$ are homeomorphisms for all α, $\beta \in \mathcal{A}$. Then one can define a topology on X by declaring the φ_α to be homeomorphisms or, in other words, that the collection

$$\{\, \varphi_\alpha^{-1}(W) \mid W \text{ open in } \mathbf{R}^n,\ \alpha \in \mathcal{A} \,\}$$

be a basis for a topology τ on X. The countability of $\mathcal{A}$ ensures that τ is second-countable, but it is not automatically Hausdorff, and this property has to be checked case-by-case. If indeed τ is Hausdorff, the collection $\{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\mathcal{A}}$ is automatically a smooth atlas for (X, τ).

Perhaps the most important example of the above is the tangent bundle of a smooth manifold. For a smooth manifold M, there is a canonical way of assembling together all of its tangent spaces at its various points. The resulting object turns out to admit a natural structure of smooth manifold of twice the dimension of M and even the structure of a vector bundle which we will discuss later.

Consider the disjoint union

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

We can view the elements of TM as equivalence classes of triples (p, a, φ), where $p \in M$, $a \in \mathbf{R}^n$ and (U, φ) is a local chart of M such that $p \in U$, and

$$(p, a, \varphi) \sim (q, b, \psi) \quad \text{if and only if } p = q \text{ and } d(\psi \circ \varphi^{-1})_{\varphi(p)}(a) = b.$$

There is a natural projection $\pi : TM \to M$ given by $\pi[p, a, \varphi] = p$, and then $\pi^{-1}(p) = T_pM$. Next, we use the above remark to introduce a topology and smooth structure on TM. Let $\{(U_\alpha, \varphi_\alpha)\}$ be a smooth atlas for M. For each α, $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha)$ is a diffeomorphism and, for each $p \in U_\alpha$, $d(\varphi_\alpha)_p : T_pU_\alpha = T_pM \to \mathbf{R}^n$ is the isomorphism mapping $[p, a, \varphi]$ to a. Set

$$\tilde{\varphi}_\alpha : \pi^{-1}(U_\alpha) \to \varphi_\alpha(U_\alpha) \times \mathbf{R}^n, \qquad [p, a, \varphi] \to (\varphi_\alpha(p), a).$$

(Equivalently, $\tilde{\varphi}_\alpha(v) = (\pi(v), d(\varphi_\alpha)_{\pi(v)}(v))$ for $v \in \pi^{-1}(U_\alpha)$.) Then $\tilde{\varphi}_\alpha$ is a bijection and $\varphi_\alpha(U_\alpha)$ is an open subset of $\mathbf{R}^{2n}$. Moreover, the maps

$$\tilde{\varphi}_\beta \circ \tilde{\varphi}_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \times \mathbf{R}^n \to \varphi_\beta(U_\alpha \cap U_\beta) \times \mathbf{R}^n$$

are given by

$$(x, a) \mapsto (\varphi_\beta \circ \varphi_\alpha^{-1}(x) \, , \, d(\varphi_\beta \circ \varphi_\alpha^{-1})_x(a)).$$

Since $\varphi_\beta \circ \varphi_\alpha^{-1}$ is a smooth diffeomorphism, we have that $d(\varphi_\beta \circ \varphi_\alpha^{-1})_x$ is a linear isomorphism and $d(\varphi_\beta \circ \varphi_\alpha^{-1})_x(a)$ is also smooth on x. It follows that $\{(\pi^{-1}(U_\alpha), \tilde{\varphi}_\alpha)\}$ defines a topology and a smooth atlas for TM so that it becomes a smooth manifold of dimension $2n$ called the *tangent bundle* of M.

Similarly, the inverses of the transpose maps of the $(d\varphi_\alpha)_p$ can be used to endow the disjoint union $T^*M := \dot{\bigcup}_{p\in M}(T_pM)^*$ of dual spaces to the tangent spaces of M with the structure of a smooth manifold of dimension $2n$, called the *cotangent bundle*. Namely, the charts have the form

$$\lambda \in \pi^{-1}(U_\alpha) \mapsto (\pi(\lambda), \left(d(\varphi_\alpha)_p^t\right)^{-1}(\lambda)) \in \varphi(U_\alpha) \times (\mathbf{R}^n)^*$$

Here $\pi : T^*M \to M$ is defined by $\pi((T_pM)^*) = \{p\}$ and $(\mathbf{R}^n)^*$ is identified with $\mathbf{R}^n$.

If $f : M \to N$ is a smooth map between smooth manifolds, we define the *differential of* f to be the map

$$df : TM \to TN$$

that restricts to $df_p : T_pM \to T_{f(p)}N$ for each $p \in M$. Using the above atlases for TM and TN, we immediately see that df is smooth.

Inverse function theorem

The proof of the following theorem just consists of unraveling the definitions and applying the inverse function theorem for smooth mappings between open subsets of $\mathbf{R}^n$.

0.2.6 Theorem (Inverse function theorem) *Let $f : M \to N$ be a smooth function between two smooth manifolds M, N, and let $p \in M$ and $q = f(p)$. Then f is a local diffeomorphism at p if and only if $df_p : T_pM \to T_qN$ is an isomorphism.*

Immersions and submanifolds

The concept of embedded submanifold that was introduced above is too strong for some purposes. There are other, weaker notions of submanifolds one of which we discuss now. We first give the following definition. A smooth map $f : M \to N$ between smooth manifolds is called an *immersion* at $p \in M$ if $df_p : T_pM \to T_{f(p)}N$ is an injective map, and f is called simply an *immersion* if it is an immersion at every point of its domain.

Let M and N be smooth manifolds such that M is a subset of N. We say that M is an *immersed submanifold of* N or simply a *submanifold of* N if the inclusion map of M into N is an immersion. Note that embedded submanifolds are automatically immersed submanifolds, but the main point behind this definition is that the topology of M can be finer than the induced topology from N. Note also that it immediately follows from this definition that if P is a smooth manifold and $f : P \to N$ is an injective immersion, then the image $f(P)$ is a submanifold of N. A smooth map $f : M \to N$ between manifolds is called an *embedding* if it is an injective immersion which is also a homeomorphism into $f(M)$ with the relative topology.

Recall that a continuous map between locally compact, Hausdorff topological spaces is called *proper* if the inverse image of a compact subset of the counter-domain is a compact subset of the domain. It is known that proper maps are closed. Also, it is clear that if the domain is compact, then every continuous map is automatically proper. An embedded submanifold M of a smooth manifold N is called *properly embedded* if the inclusion map is proper. Now the following proposition is a simple remark.

0.2.7 Proposition *If $f : M \to N$ is an injective immersion which is also a proper map, then the image $f(M)$ is a properly embedded submanifold of N.*

As an application of the inverse function theorem, it is not difficult to see that any immersion $f : M \to N$, where $\dim M = n$, $\dim N = n + k$, can be locally represented via appropriate charts as the standard inclusion $\mathbf{R}^n \to \mathbf{R}^{n+k}$ In particular, it is locally an embedding. This result will be particularly useful in geometry when dealing with local properties of an isometric immersion. It also follows from the local form of an immersion that *the image of an embedding is an embedded submanifold*.

0.2.8 Example Take the 2-torus $T^2 = S^1 \times S^1$ viewed as a submanifold of $\mathbf{R}^2 \times \mathbf{R}^2 = \mathbf{R}^4$ and consider the map

$$f : \mathbf{R} \to T^2, \qquad f(t) = (\cos at, \sin at, \cos bt, \sin bt),$$

where a, b are non-zero real numbers. Since $f'(t)$ never vanishes, this map is an immersion. If b/a is rational, f is periodic and f induces an embeddeding of S^1 into T^2. If b/a is an irrational number, then $f(\mathbf{R})$ is *not* an embedded submanifold of T^2. In fact, the assumption on b/a implies that $f(\mathbf{R})$ is a dense subset of T^2, but an embedded submanifold of some other manifold is always locally closed. ⋆

Submersions and inverse images

Submanifolds can also be defined by equations togetehr with some nondegeracy conditions. In order to explain this point, we introduce the following definition. A smooth map $f : M \to N$ between manifolds is called a *submersion* at $p \in M$ if $df_p : T_pM \to T_{f(p)}N$ is a surjective map, and f is called simply a *submersion* if it is a submersion at every point of its domain.

As an application of the inverse function theorem, it is not difficult to see that any submersion $f : M \to N$, where $\dim M = n + k$, $\dim N = n$, can be locally represented via appropriate charts as the standard projection $\mathbf{R}^{n+k} \to \mathbf{R}^n$. It follows that each level set of f admits the structure of an embedded submanifold of dimension k.

0.2.9 Examples (a) Let A be a non-degenerate real symmetric matrix of order $n + 1$ and define $f : \mathbf{R}^{n+1} \to \mathbf{R}$ by $f(p) = \langle Ap, p\rangle$ where $\langle , \rangle$ is the standard Euclidean inner product. Then $df_p : \mathbf{R}^{n+1} \to \mathbf{R}$ is given by $df_p(v) = 2\langle Ap, v\rangle$, so it is surjective if $p \neq 0$. It follows that f is a submersion on $\mathbf{R}^{n+1} \setminus \{0\}$ and $f^{-1}(r)$ for $r \in \mathbf{R}$ is an embedded submanifold of $\mathbf{R}^{n+1}$ of dimension n if it is nonempty. In particular, by taking A to be the identity matrix we get a manifold structure for S^n which coincides with the one previously constructed.

(b) Denote by V the vector space of real symmetric matrices of order n, and define $f : GL(n,\mathbf{R}) \to V$ by $f(A) = AA^t$. We first claim that f is a submersion at the identity matrix I. One easily computes that

$$df_I(B) = \lim_{h\to 0} \frac{f(I+hB)-f(I)}{h} = B + B^t,$$

where $B \in T_I GL(n,\mathbf{R}) = M(n,\mathbf{R})$. Now, given $C \in V$, df_I maps $\frac{1}{2}C$ to C, so this checks the claim. We next check that f is a submersion at any $D \in f^{-1}(I)$. Note that $DD^t = I$ implies that $f(AD) = f(A)$. This means that $f = f \circ R_D$, where $R_D : GL(n,\mathbf{R}) \to GL(n,\mathbf{R})$ is the map that multiplies on the right by D. We have that R_D is a diffeomorphism of $GL(n,\mathbf{R})$ whose inverse is plainly given by $R_{D^{-1}}$. Therefore $d(R_D)_I$ is an isomorphism, so the chain rule $df_I = df_D \circ d(R_D)_I$ yields that df_D is surjective, as desired. Now $f^{-1}(I) = \{ A \in GL(n,\mathbf{R}) \,|\, AA^t = I \}$ is an embedded submanifold of $GL(n,\mathbf{R})$ of dimension

$$\dim GL(n,\mathbf{R}) - \dim V = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

Note that $f^{-1}(I)$ is a group with respect to the multiplication of matrices; it is called the *orthogonal group* of order n and is usually denoted by $O(n)$. ⋆

Smooth coverings

In this subsection, we summarize some properties of covering spaces in the context of smooth manifolds. Recall that a (topological) *covering* of a space X is another space $\tilde{X}$ with a continuous map $\pi : \tilde{X} \to X$ such that X is a union of evenly covered open set, where a connected open subset U of X is called *evenly covered* if

$$\pi^{-1}U = \cup_{i\in I}\tilde{U}_i \tag{0.2.10}$$

is a disjoint union of open sets $\tilde{U}_i$ of $\tilde{X}$, each of which is mapped homeomorphically onto U under π. In particular, the fibers of π are discrete subsets of $\tilde{X}$. It also follows from the definition that $\tilde{X}$ has the Hausdorff property if X does. Further it is usual, as we shall do, to require that X and $\tilde{X}$ be connected, and then the index set I can be taken the same for all evenly covered open sets.

0.2.11 Examples (a) $\pi : \mathbf{R} \to S^1$, $\pi(t) = e^{it}$ is a covering.

(b) $\pi : S^1 \to S^1$, $\pi(z) = z^n$ is a covering for any nonzero integer n.

(c) $\pi : (0, 3\pi) \to S^1$, $\pi(t) = e^{it}$ is a local homemeomorphism which is not a covering, since $1 \in S^1$ does not admit evenly covered neighborhoods. ⋆

Covering spaces are closely tied with fundamental groups. The *fundamental group* $\pi_1(X, x_0)$ of a topological space X with basepoint x_0 is defined as follows. As a set, it consists of the homotopy classes of continuous loops based at x_0. The concatenation of such loops is compatible with the equivalence relation given by homotopy, so it induces a group operation on $\pi_1(X, x_0)$ making it into a group. If X is arcwise connected, the isomorphism class of the fundamental group is independent of the choice of basepoint (indeed for x_0, $x_1 \in X$ and c a continuous path from x_0 to x_1, conjugation by c^{-1} induces an isomorphism from $\pi_1(X, x_0)$ and $\pi_1(X, x_1)$) and thus is sometimes denoted by $\pi_1(X)$. Finally, a continuous map $f : X \to Y$ between topological spaces with $f(x_0) = y_0$ induces a homomorphism $f_\# : \pi_1(X, x_0) \to \pi_1(Y, y_0)$ so that the assignment $(X, x_0) \to \pi_1(X, x_0)$ is functorial. Of course the fundamental group is trivial if and only if the space is simply-connected.

Being locally Euclidean, a smooth manifold is locally arcwise connected and locally simply-connected. A connected space X with such local connectivity properties admits a simply-connected covering space, which is unique up to isomorphism; an isomorphism between coverings $\pi_1 : \tilde{X}_1 \to X$ and $\pi_2 : \tilde{X}_2 \to X$ is a homeomorphism $f : \tilde{X}_1 \to \tilde{X}_2$ such that $\pi_2 \circ f = \pi_1$. More generally, there exists a bijective correspondance between isomorphism classes of coverings $\pi : (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ and subgroups of $\pi_1(X, x_0)$ given by $(\tilde{X}, \tilde{x}_0) \mapsto \pi_\#(\pi_1(\tilde{X}, \tilde{x}_0))$; moreover, a change of basepoint in $\tilde{X}$ corresponds to passing to a conjugate subgroup $\pi_1(X, x_0)$.

Suppose $\pi : \tilde{M} \to M$ is a covering where M is a smooth manifold. Then there is a natural structure of smooth manifold on $\tilde{M}$ such that the projection π is smooth. In fact, for every chart (U, π) of M where U is evenly covered as in (0.2.10), take a chart $(\tilde{U}_i, \varphi \circ \pi|_{\tilde{U}_i})$ for $\tilde{M}$. This gives an atlas of $\tilde{M}$, which is smooth because for another chart (V, ψ) of M, V evenly covered by $\cup_{i \in I} \tilde{V}_i$ and $\tilde{U}_i \cap \tilde{V}_j \neq \varnothing$ for some i, $j \in I$, we have that the transition map

$$(\psi \circ \pi|_{\tilde{V}_j})(\varphi \circ \pi|_{\tilde{U}_i})^{-1} = \psi \circ \varphi^{-1}$$

is smooth. We already know that $\tilde{M}$ is a Hausdorff space. It is possible to choose a countable basis of connected open sets for M which are evenly covered. The connected components of the preimages under π of the elements of this basis form a basis of connected open sets for $\tilde{M}$, which is countable as long as the index set I is countable, but this follows from the countability of the fundamental group $\pi_1(M)$■1■. Now, around any point in $\tilde{M}$, π admits a local representation as the identity, so it is a local diffeomorphism. Note that we have indeed proved more: M can be covered by evenly covered neighborhoods U such that the restriction of π to a connected component of $\pi^{-1}U$ is a diffeomorphism onto U. This is the definition of a *smooth covering*. Note that a topologic covering whose covering map is smooth need not be a smooth covering (e.g. $\pi : \mathbf{R} \to \mathbf{R}$, $\pi(x) = x^3$).

Next, we can formulate basic results in covering theory for a smooth covering $\pi : \tilde{M} \to M$ of a smooth manifold M. Fix basepoints $\tilde{p} \in \tilde{M}$, $p \in M$ such that $\pi(\tilde{p}) = p$. We say that a map $f : N \to M$ admits a lifting if there exists a map $\tilde{f} : N \to \tilde{M}$ such that $\pi \circ \tilde{f} = f$.

0.2.12 Theorem (Lifting criterion) *Let $q \in f^{-1}(p)$. A smooth map $f : N \to M$ admits a smooth lifting $\tilde{f} : N \to \tilde{M}$ with $\tilde{f}(q) = \tilde{p}$ if and only if $f_\#(\pi_1(N, q)) \subset \pi_\#(\pi_1(\tilde{M}, \tilde{p}))$. In that case, if N is connected, the lifting is unique.*

Taking $f : N \to M$ to be the universal covering of M in Theorem 0.2.12 shows that the universal covering of M covers any other covering of M and hence justifies its name.

For a topological covering $\pi : \tilde{X} \to X$, a *deck transformation* or *covering transformation* is an isomorphism $\tilde{X} \to \tilde{X}$, namely, a homeomorphism $f : \tilde{X} \to \tilde{X}$ such that $\pi \circ f = \pi$. The deck transformations form a group under composition. It follows from uniqueness of liftings that a deck transformation is uniquely determined by its action on one point. In particular, the only deck transformation admitting fixed points is the identity. Since a smooth covering map $\pi : \tilde{M} \to M$ is a local diffeomorphism, in this case the equation $\pi \circ f = \pi$ implies that deck transformations are diffeomorphisms of $\tilde{M}$.

An action of a (discrete) group on a topological space (resp. smooth manifold) is a homomorphism from the group to the group of homeomorphisms (resp. diffeomorphisms) of the space (resp. manifold). For a smooth manifold M, we now recall the canonical action of $\pi_1(M, p)$ on its universal covering $\tilde{M}$ by deck transformations. First we remark that by the lifting criterion, given $q \in M$ and $\tilde{q}_1$, $\tilde{q}_2 \in \pi^{-1}(q)$, there is a unique deck transformation mapping $\tilde{q}_1$ to $\tilde{q}_2$. Now let γ be a continuous loop in M based at p representing an element $[\gamma] \in \pi_1(M, p)$. By the remark, it

■1■Ref?

suffices to describe the action of $[\gamma]$ on a point $\tilde{p} \in \pi^{-1}(p)$, which goes as follows: lift γ uniquely to a path $\tilde{\gamma}$ starting at $\tilde{p}$; then $[\gamma] \cdot \tilde{p}$ is by definition the endpoint of $\tilde{\gamma}$, which sits in the fiber $\pi^{-1}(p)$. The definiton independs of the choice made, namely, if we change γ to a homotopic curve, we get the same result. This follows from Theorem 0.2.12 applied to the homotopy, as it is defined on a square and a square is simply-connected. Since $\pi : \tilde{M} \to M$ is the universal covering, every deck transformation is obtained in this way from an element of $\pi_1(M, p)$.

As action of a (discrete) group Γ on a topological space X is called *free* if no nontrivial element of Γ has fixed points, and it is called *proper* if any two points x, $y \in X$ admit open neighborhoods $U \ni x$, $V \ni y$ such that $\{ \gamma \in \Gamma \mid \gamma U \cap V \neq \varnothing \}$ is finite. The action of $\pi_1(M, p)$ on the universal covering $\tilde{M}$ by deck transformations has both properties. In fact, we have already remarked it is free. To check properness, let $\tilde{p}$, $\tilde{q} \in \tilde{M}$. If these points lie in the same orbit of $\pi_1(M, p)$ or, equivalently, the same fiber of π, the required neighborhoods are the connected components of $\pi^{-1}(U)$ containing $\tilde{p}$ and $\tilde{q}$, resp., where U is an evenly covered neighborhood of $\pi(\tilde{p}) = \pi(\tilde{q})$. On the other hand, if $\pi(\tilde{p}) =: p \neq q := \pi(\tilde{p})$, we use the Hausdorff property of M to find disjoint evenly covered neighborhoods $U \ni p$, $V \ni q$ and then it is clear that the connected component of $\pi^{-1}(U)$ containing $\tilde{p}$ and the connected component of $\pi^{-1}(V)$ containing $\tilde{q}$ do the job.

Conversely, we have:

0.2.13 Theorem *If the group Γ acts freely and properly on a smooth manifold $\tilde{M}$, then the quotient space $M = \Gamma \backslash \tilde{M}$ endowed with the quotient topology admits a unique structure of smooth manifold such that the projection $\pi : \tilde{M} \to M$ is a smooth covering.*

Proof. The action of Γ on $\tilde{M}$ determines a partition into equivalence classes or *orbits*, namely $\tilde{p} \sim \tilde{q}$ if and only if $\tilde{q} = \gamma \tilde{p}$ for some $\gamma \in \Gamma$. The orbit through $\tilde{p}$ is denoted $\Gamma(\tilde{p})$. The quotient space $\Gamma \backslash \tilde{M}$ is also called *orbit space.*

The quotient topology is defined by the condition that $U \subset M$ is open if and only if $\pi^{-1}(U)$ is open in $\tilde{M}$. In particular, for an open set $\tilde{U} \subset \tilde{M}$ we have $\pi^{-1}(\pi(\tilde{U})) = \cup_{\gamma \in \Gamma} \gamma(\tilde{U})$, a union of open sets, showing that $\pi(\tilde{U})$ is open and proving that π is an open map. In particular, π maps a countable basis of open sets in $\tilde{M}$ to a countable basis of open sets in M.

The covering property follows from the fact that Γ is proper. In fact, let $\tilde{p} \in \tilde{M}$. From the definition of properness, we can choose a neighborhood $\tilde{U} \ni \tilde{p}$ such that $\{ \gamma \in \Gamma \mid \gamma \tilde{U} \cap \tilde{U} \neq \varnothing \}$ is finite. Using the Hausdorff property of $\tilde{M}$ and the freeness of Γ, we can shrink $\tilde{U}$ so that this set contains the identity only. Now the map π identifies all disjoint homeomorphic open sets γU for $\gamma \in \Gamma$ to a single open set $\pi(U)$ in M, which is then evenly covered.

The Hausdorff property of M also follows from properness of Γ. Indeed, let p, $q \in M$, $p \neq q$. Choose $\tilde{p} \in \pi^{-1}(p)$, $\tilde{q} \in \pi^{-1}(q)$ and neighborhoods $\tilde{U} \ni \tilde{p}$, $\tilde{V} \ni \tilde{q}$ such that $\{ \gamma \in \Gamma \mid \gamma \tilde{U} \cap \tilde{V} \neq \varnothing \}$ is finite. Note that $\tilde{q} \notin \Gamma(\tilde{p})$, so by the Hausdorff property for $\tilde{M}$, we can shrink $\tilde{U}$ so that this set becomes empty. Since π is open, $U := \pi(\tilde{U})$ and $V := \pi(\tilde{V})$ are now disjoint neighborhoods of p and q, respectively.

Finally, we construct a smooth atlas for M. Let $p \in M$ and choose an evenly covered neighborhood $U \ni p$. Write $\pi^{-1} U = \cup_{i \in I} \tilde{U}_i$ as in (0.2.10). By shrinking U we can ensure that $\tilde{U}_i$ is the domain of a local chart $(\tilde{U}_i, \tilde{\varphi}_i)$ of $\tilde{M}$. Now $\varphi_i := \tilde{\varphi}_i \circ (\pi|_{\tilde{U}_i})^{-1} : U \to \mathbf{R}^n$ defines a homeomorphism onto the open set $\tilde{\varphi}_i(\tilde{U}_i)$ and thus a local chart (U, φ_i) of M. The domains of such charts cover M and it remains only to check that the transition maps are smooth. So let V be another evenly covered neighborhood of p with $\pi^{-1} V = \cup_{j \in I} \tilde{V}_j$ and associated local chart $\psi_j := \tilde{\psi}_j \circ (\pi|_{\tilde{V}_j})^{-1} : U \to \mathbf{R}^n$ where $(\tilde{V}_j, \tilde{\psi}_j)$ is a local chart of $\tilde{M}$. Then

$$\psi_j \circ \varphi_i^{-1} = \tilde{\psi}_j \circ (\pi|_{\tilde{V}_j})^{-1} \circ \pi \circ \tilde{\varphi}_i^{-1} \tag{0.2.14}$$

However, $(\pi|_{\tilde{V}_j})^{-1} \circ \pi$ is realized by a unique element $\gamma \in \Gamma$ in a neighborhood of $\tilde{p}_i = \pi|_{\tilde{U}_i}^{-1}(p)$. Since Γ acts by diffeomorphisms, this shows that the transtion map (0.2.14) is smooth and finishes the proof. □

0.3 Vector fields

A *vector field* X on a smooth manifold M is an assignment of a vector $X(p)$ in each T_pM. Sometimes we write X_p for $X(p)$. Vector fields are the infinitesimal objects associated to diffeomorphisms in the following sense. Let $\varphi_t : M \to M$ be a diffeomorphism such that the curve $t \mapsto \varphi_t(p)$ is smooth for each p. Then $X_p := \frac{d}{dt}\big|_{t=0}\varphi_t(p)$ defines a vector field on M. Conversely, one can integrate smooth vector fields to obtain diffeomorphisms. Actually, this is the extension of ODE theory to smooth manifolds that we shall recall below.

We need the notion of smoothness for vector fields. Recall that TM is a smooth manifold, so a vector field $X : M \to TM$ is called *smooth* simply if this map is smooth.

For practical purposes, we reformulate this notion. Let X be an arbitrary vector field on M. Given a smooth function f on an open subset U of M, the directional derivative $X(f) : U \to \mathbf{R}$ is defined to be the function $p \in U \mapsto X_p(f)$. Further, if $(x_1, \ldots, x_n)$ is a coordinate system on U, we have already seen that $\{\frac{\partial}{\partial x_1}|_p, \ldots, \frac{\partial}{\partial x_n}|_p\}$ is a basis of T_pM for $p \in U$. It follows that there are functions $a_i : U \to \mathbf{R}$ such that

$$X|_U = \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i}. \tag{0.3.1}$$

0.3.2 Proposition *Let X be a vector field on M. Then the following assertions are equivalent:*

a. X is smooth.

b. For every coordinate system $(U, (x_1, \ldots, x_n))$ of M, the functions a_i defined by (0.3.1) are smooth.

c. For every open set V of M and smooth map $f : V \to \mathbf{R}$, the function $X(f) : V \to \mathbf{R}$ is smooth.

Since $a_i = X(x_i)$ in (0.3.1), we have

0.3.3 Scholium *If X is a smooth vector field on M and $X(f) = 0$ for every smooth function, then $X = 0$.*

We now come to the integration of smooth vector fields. Let X be a smooth vector field on M An *integral curve* of X is a smooth curve $\gamma : I \to M$, where I is an open interval, such that

$$\dot{\gamma}(t) = X(\gamma(t))$$

for all $t \in I$. We write this equation in local coordinates. Suppose X has the form (0.3.1), $\gamma_i = x_i \circ \gamma$ and $\tilde{a}_i = a_i \circ \varphi^{-1}$. Then γ is an integral curve of X in $\gamma^{-1}(U)$ if and only if

$$\frac{d\gamma_i}{dr}\bigg|_t = \tilde{a}_i(\gamma_1(t), \ldots, \gamma_n(t)) \tag{0.3.4}$$

for $i = 1, \ldots, n$ and $t \in \gamma^{-1}(U)$. Equation (0.3.4) is a system of first order ordinary differential equations for which existence and uniqueness theorems are known. These, translated into manifold terminology yield local existence and uniqueness of integral curves for smooth vector fields. Moreover, one can cover M by domains of local charts and piece together the locally defined integral curves of X to obtain, for any given point $p \in M$, a *maximal* integral curve γ_p of X through p defined on a possibly infinite interval $(a(p), b(p))$.

Even more interesting is to reverse the rôles of p and t by setting

$$\varphi_t(p) := \gamma_p(t)$$

for all p such that $t \in (a(p), b(p))$.

The smooth dependence of solutions of ODE on the initial conditions implies that for every $p \in M$, there exists an open neighborhood V of p and $\epsilon > 0$ such that the map

$$(-\epsilon, \epsilon) \times V \to M, \quad (t, p) \mapsto \varphi_t(p)$$

is well defined and smooth. Glueing integral curves one checks that

$$\varphi_{s+t} = \varphi_s \circ \varphi_t \tag{0.3.5}$$

whenever both hand sides are defined. Obviously φ_0 is the identity, so φ_t is a diffeomorphism defined on some open subset of M with inverse φ_{-t}. The collection $\{\varphi_t\}$ is called the *flow* of X. Owing to property (0.3.5), the flow of X is also called the *one-parameter local group* of locally defined diffeomorphisms generated by X, and X is called the *infinitesimal generator* of $\{\varphi_t\}$. If φ_t is defined for all $t \in \mathbf{R}$, the vector field X is called *complete*. This is equivalent to requiring that the maximal integral curves of X be defined on the entire $\mathbf{R}$, or yet, that the domain of each φ_t be M. In this case we refer to $\{\varphi_t\}$ as the *one-parameter group* of diffeomorphisms of M generated by X.

0.3.6 Examples (a) Take $M = \mathbf{R}^2$ and $X = \frac{\partial}{\partial x_1}$. Then X is complete and $\varphi_t(x_1, x_2) = (x_1 + t, x_2)$ for $(x_1, x_2) \in \mathbf{R}^2$. Note that if we replace $\mathbf{R}^2$ by the punctured plane $\mathbf{R}^2 \setminus \{(0, 0)\}$, the domains of φ_t become proper subsets of M.

(b) Consider the smooth vector field on $\mathbf{R}^{2n}$ defined by

$$X(x_1, \ldots, x_{2n}) = -x_2 \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2} + \cdots - x_{2n} \frac{\partial}{\partial x_{2n-1}} + x_{2n-1} \frac{\partial}{\partial x_{2n}}.$$

The flow of X is given the linear map

$$\varphi_t \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{2n-1} \\ x_{2n} \end{pmatrix} = \begin{pmatrix} R_t & & \\ & \ddots & \\ & & R_t \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{2n-1} \\ x_{2n} \end{pmatrix}$$

where R_t is the 2×2 block

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

It is clear that X restricts to a smooth vector field $\bar{X}$ on S^{2n-1}. The flow of $\bar{X}$ is of course the restriction of φ_t to S^{2n-1}. X and $\bar{X}$ are complete vector fields.

(c) Take $M = \mathbf{R}$ and $X(x) = x^2 \frac{\partial}{\partial x}$. Solving the ODE we find $\varphi_t(x) = \frac{x}{1-tx}$. It follows that the domain of φ_t is $(-\infty, \frac{1}{t})$ if $t > 0$ and $(\frac{1}{t}, +\infty)$ if $t < 0$. ⋆

Lie bracket

If X is a smooth vector field on M and $f : M \to \mathbf{R}$ is a smooth function, the directional derivative $X(f) : M \to \mathbf{R}$ is also smooth and so it makes sense to derivate it again as in $Y(X(f))$ where Y is another smooth vector field on M. For instance, in a local chart $(U, \varphi = (x_1, \ldots, x_n))$, we have the first order partial derivative

$$\frac{\partial}{\partial x_i}\Big|_p (f) = \frac{\partial f}{\partial x_i}\Big|_p$$

and the second order partial derivative

$$\left(\frac{\partial}{\partial x_j}\right)_p \left(\frac{\partial}{\partial x_i}(f)\right) = \frac{\partial^2 f}{\partial x_j \partial x_i}\Big|_p$$

and it follows from Schwarz theorem on the commutativity of mixed partial derivatives of smooth functions on $\mathbf{R}^n$ that

$$\text{(0.3.7)} \qquad \frac{\partial^2 f}{\partial x_j \partial x_i}\Big|_p = \frac{\partial^2 (f \circ \varphi^{-1})}{\partial r_j \partial r_i}\Big|_p = \frac{\partial^2 (f \circ \varphi^{-1})}{\partial r_i \partial r_j}\Big|_p = \frac{\partial^2 f}{\partial x_i \partial x_j}\Big|_p,$$

where $\mathrm{id} = (r_1, \ldots, r_n)$ denote the canonical coordinates on $\mathbf{R}^n$.

On the other hand, for general smooth vector fields X, Y on M the second derivative depends on the order of the vector fields and the failure of the commutativity is measured by the *commutator* or *Lie bracket*

$$\text{(0.3.8)} \qquad [X, Y](f) = X(Y(f)) - Y(X(f))$$

for every smooth function $f : M \to \mathbf{R}$. We say that X, Y *commute* if $[X, Y] = 0$. It turns out that formula (0.3.8) defines a smooth vector field on M! Indeed, Scholium 0.3.3 says that such a vector field is unique, if it exists. In order to prove existence, consider a coordinate system $(U, (x_1, \ldots, x_n))$. Then we can write

$$X|_U = \sum_{i=1}^n a_i \frac{\partial}{\partial x_i} \quad \text{and} \quad Y|_U = \sum_{j=1}^n b_j \frac{\partial}{\partial x_j}$$

for a_i, $b_j \in C^\infty(U)$. If $[X, Y]$ exists, we must have

$$\text{(0.3.9)} \qquad [X, Y]|_U = \sum_{i=1}^n \left(a_i \frac{\partial b_j}{\partial x_i} - b_i \frac{\partial a_j}{\partial x_i}\right) \frac{\partial}{\partial x_j},$$

because the coefficients of $[X, Y]|_U$ in the local frame $\{\frac{\partial}{\partial x_j}\}_{j=1}^n$ must be given by $[X, Y](x_j) = X(Y(x_j)) - Y(X(x_j))$. We can use formula (0.3.9) as the definition of a vector field on U; note that such a vector field is smooth and satisfies property (0.3.8) for functions in $C^\infty(U)$. We finally define $[X, Y]$ globally by covering M with domains of local charts: on the overlap of two charts, the different definitions coming from the two charts must agree by the above uniqueness result; it follows that $[X, Y]$ is well defined.

0.3.10 Examples (a) Schwarz theorem (0.3.7) now means $[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0$ for coordinate vector fields associated to a local chart.

(b) Let $X = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial z}$, $Y = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial z}$, $Z = \frac{\partial}{\partial z}$ be smooth vector fields on $\mathbf{R}^3$. Then $[X, Y] = Z$, $[Z, X] = [Z, Y] = 0$. ⋆

The proof of the following proposition only uses (0.3.8).

0.3.11 Proposition *Let X, Y and Z be smooth vector fields on M. Then*
a. $[Y,X] = -[X,Y]$.
b. If f, $g \in C^\infty(M)$, then

$$[fX, gY] = fg[X,Y] + f(Xg)Y - g(Yf)X.$$

c. $[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0$. *(Jacobi identity)*

Let $f : M \to N$ be a diffeomorphism. For every smooth vector field X on M, the formula $df \circ X \circ f^{-1}$ defines a smooth vector field on N which we denote by f_*X. If the flow is $\{\varphi_t\}$, then the flow of f_*X is $f \circ \varphi_t \circ f^{-1}$. More generally, if $f : M \to N$ is a smooth map which needs not be a diffeomorphism, smooth vector fields X on M and Y on N are called *f-related* if $df \circ X = Y \circ f$. The proof of the next propostion is an easy application of (0.3.8).

0.3.12 Proposition *Let $f : M \to M'$ be smooth. Let X, Y be smooth vector fields on M, and let X', Y' be smooth vector fields on M'. If X and X' are f-related and Y and Y' are f-related, then also $[X,Y]$ and $[X',Y']$ are f-related.*

What is the relation between flows and Lie brackets? In order to discuss that, let X, Y be smooth vector fields on M with corresponding flows $\{\varphi_t\}$, $\{\psi_s\}$. Fix $p \in M$ and a smooth function f defined on a neighborhood of p. We have

$$\begin{aligned}
[X,Y]_p(f) &= X_p(Yf) - Y_p(Xf) \\
&= \frac{d}{dt}\Big|_{t=0}(Yf)(\varphi_t(p)) - \frac{d}{ds}\Big|_{s=0}(Xf)(\psi_s(p)) \\
&= \frac{\partial^2}{\partial s \partial t}\Big|_{(0,0)} f(\psi_s(\varphi_t(p))) - \frac{\partial^2}{\partial t \partial s}\Big|_{(0,0)} f(\varphi_t(\psi_s(p))) \\
&= \frac{\partial^2}{\partial t \partial s}\Big|_{(0,0)} f(\varphi_{-t}(\psi_s(\varphi_t(p)))) \\
&= \frac{d}{dt}\Big|_{t=0}\left(\frac{d}{ds}\Big|_{s=0} f(\varphi_{-t} \circ \psi_s \circ \varphi_t(p))\right) \\
&= \frac{d}{dt}\Big|_{t=0} \left((\varphi_{-t})_*Y\right)_p(f)
\end{aligned}$$

Note that $t \mapsto ((\varphi_{-t})_*Y)_p$ is a smooth curve in T_pM. Its tangent vector at $t = 0$ is called the *Lie derivative* of Y with respect to X at p, denoted by $(L_XY)_p$, and this defines the Lie derivative L_XY as a smooth vector field on M. The above calculation shows that $L_XY = [X,Y]$.

0.3.13 Proposition *X and Y commute if and only if their corresponding flows $\{\varphi_t\}$, $\{\psi_s\}$ commute.*

Proof. $[X,Y] = 0$ if and only if $0 = \frac{d}{dt}\Big|_{t=0}(\varphi_{-t})_*Y$. Since $\{\varphi_t\}$ is a one-parameter group, this is equivalent to $(\varphi_{-t})_*Y = Y$ for all t. However the flow of $(\varphi_{-t})_*Y$ is $\{\varphi_{-t}\psi_s\varphi_t\}$, so this means $\varphi_{-t}\psi_s\varphi_t = \psi_s$. □

We know that, for a local chart (U,φ), the set of coordinate vector fields $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\}$ is linearly independent at every point of U and the $\frac{\partial}{\partial x_i}$ pairwise commute. It turns out these two conditions locally characterize coordinate vector fields. Namely, we call a set $\{X_1, \ldots, X_n\}$ of smooth vector fields defined on an open set V of M a *local frame* if it is linearly independent at every point of V.

0.3.14 Proposition *Let $\{X_1, \ldots, X_n\}$ be a local frame on V such that $[X_i, X_j] = 0$ for all i, $j = 1, \ldots, n$. Then for every $p \in V$ there exists an open neighborhood U of p in V and a local chart (U, φ) whose coordinate vector fields are exactly the X_i.*

Proof. Let $\{\varphi_t^i\}$ be the flow of X_i and put $F(t_1, \ldots, t_n) := \varphi_{t_1}^1 \circ \cdots \circ \varphi_{t_n}^n(p)$, defined on a neighborhood of 0 in $\mathbf{R}^n$. Then $dF_0(e_i) = X_i(p)$ for all i, so F is a local diffeomorphism at 0 by the inverse function theorem. The local inverse F^{-1} defines a local chart around p. Finally, $\frac{\partial}{\partial x_i} = X_i$ by Proposition 0.3.13. □

0.4 Lie groups

Lie groups comprise a very important class of examples of smooth manifolds. At the same time, they are used to model transformation groups of smooth manifolds.

A *Lie group* G is a smooth manifold endowed with a group structure such that the group operations are smooth. More concretely, the multiplication map $\mu : G \times G \to G$ and the inversion map $\iota : G \to G$ are required to be smooth.

0.4.1 Examples (a) The Euclidean space $\mathbf{R}^n$ with its additive vector space structure is a Lie group. Since the multiplication is commutative, this is an example of a *Abelian* (or *commutative*) Lie group.

(b) The multiplicative group of nonzero complex numbers $\mathbf{C}^\times$. The subgroup of unit complex numbers is also a Lie group, and as a smooth manifold it is diffeomorphic to the circle S^1.

(c) If G and H are Lie groups, the direct product group structure turns the product manifold $G \times H$ into a Lie group.

(d) It follows from (b) and (c) that the *n-torus* $T^n = S^1 \times \cdots \times S^1$ (n times) is a Lie group. Of course, T^n is a compact connected Abelian Lie group. Conversely, we will see in Theorem 0.4.13 that every compact connected Abelian Lie group is an n-torus.

(e) If G is a Lie group, the connected component of the identity of G, denoted by G°, is also a Lie group. Indeed, G° is open in G, so it inherits a smooth structure from G just by restricting the local charts. Since $\mu(G^\circ \times G^\circ)$ is connected and $\mu(1,1) = 1$, we must have $\mu(G^\circ \times G^\circ) \subset G^\circ$. Similarly, $\iota(G^\circ) \subset G^\circ$. Since $G^\circ \subset G$ is an open submanifold, it follows that the group operations restricted to G° are smooth.

(f) Any finite or countable group endowed with the discrete topology becomes a 0-dimensional Lie group. Such examples are called *discrete Lie groups*.

(g) We now turn to some of the classical matrix groups. The *real general linear group of order* n, which is denoted by $\mathbf{GL}(n, \mathbf{R})$, is the group consisting of all nonsingular $n \times n$ real matrices. Denote by $M(n, \mathbf{R})$ the vector space of all $n \times n$ real matrices and consider the determinant function $\det : M(n, \mathbf{R}) \to \mathbf{R}$. Since $GL(n, \mathbf{R})$ consists precisely of the matrices in $M(n, \mathbf{R})$ with nonzero determinant, we see that $\mathbf{GL}(n, \mathbf{R})$ is open in $M(n, \mathbf{R})$ and thus inherits the structure of a smooth manifold. In the coordinates provided by the canonical identification $M(n, \mathbf{R}) \cong \mathbf{R}^{n^2}$, the group operations of $GL(n, \mathbf{R})$ are expressed by rational functions and are thus smooth. Note that $\dim GL(n, \mathbf{R}) = n^2$. Similarly, one defines the *complex general linear group of order* n, which is denoted by $\mathbf{GL}(n, \mathbf{C})$, as the group consisting of all nonsingular $n \times n$ complex matrices. Note that $\dim GL(n, \mathbf{C}) = 2n^2$.

We have already encoubtered the orthogonal group $O(n)$ as a closed embedded submanifold of $GL(n,\mathbf{R})$ in 0.2.9. Since $O(n)$ is an embedded submanifold, it follows from Theorem **??** that the group operations of $O(n)$ are smooth, and hence $O(n)$ is a Lie group. ⋆

At this juncture, it is convenient to introduce another object. A *(real, complex) Lie algebra* is a (real, complex) vector space $\mathfrak{g}$ endowed with a bilinear operation

$$[\cdot,\cdot]:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$$

satisfying:
(a) $[Y,X]=-[X,Y]$ (skew-symmetry); and
(b) $[[X,Y],Z]+[[Y,Z],X]+[[Z,X],Y]=0$ (Jacobi identity); where X, Y, $Z\in\mathfrak{g}$.
Of course, a Lie algebra is a nonassociative, in general noncommutative algebra in which the commutative and associative properties have been replaced by (a) and (b) above. It is clear that (a) is equivalent to having $[X,X]=0$ for all $X\in\mathfrak{g}$, and that identity (b) only imposes additional restrictions if X, Y, Z are linearly independent.

0.4.2 Examples (a) Let M be a smooth manifold and consider the infinite-dimensional real vector space $\mathfrak{X}(M)$ of all smooth vector fields on M. It follows from Proposition 0.3.11 that $\mathfrak{X}(M)$ equipped with the Lie bracket is an infinite-dimensional Lie algebra.

(b) Let V be any vector space and take $[\cdot,\cdot]$ to be the zero bilinear form. Then V becomes a so called *Abelian* Lie algebra.

(c) Let A be any real associative algebra and set $[a,b]=ab-ba$ for a, $b\in A$. It is easy to see that A becomes a Lie algebra. An important instance of this situation is $A=M(n,\mathbf{R})$; the associated Lie algebra is sometimes denoted by $\mathfrak{gl}(n,\mathbf{R})$.

(d) The subset of $\mathfrak{gl}(n,\mathbf{R})$ consisting of skew-symmetric matrices is closed under the Lie bracket and hence is a Lie algebra itself, denoted by $\mathfrak{so}(n)$.

(e) The cross-product $\times$ on $\mathbf{R}^3$ is easily seen to define a Lie algebra structure.

(f) If V is a two-dimensional vector space and X, $Y\in V$ are linearly independent, the conditions $[X,X]=[Y,Y]=0$, $[X,Y]=X$ define a Lie algebra structure on V.

(g) If V is a three-dimensional vector space spanned by X, Y, $Z\in V$, the conditions $[X,Y]=Z$, $[Z,X]=[Z,Y]=0$ define a Lie algebra structure on V, called the *(3-dimensional) Heisenberg algebra*. It can be realized as a Lie algebra of smooth vector fields on $\mathbf{R}^3$ as in example 0.3.10(b). ⋆

One of the most essential features of Lie groups is the existence of translations. Let G be a Lie group. The *left translation* defined by $g\in G$ is the map $L_g:G\to G$, $L_g(x)=gx$. It is a diffeomorphism of G, its inverse being given by $L_{g^{-1}}$. Similarly, the *right translation* defined by $g\in G$ is the map $R_g:G\to G$, $R_g(x)=xg$. It is also a diffeomorphism of G, and its inverse is given by $R_{g^{-1}}$.

The translations in G allow us to consider invariant tensors, the most important case being that of vector fields. A vector field X on G is called *left-invariant* if $d(L_g)_x(X_x)=X_{gx}$ for every g, $x\in X$. This condition is simply $dL_g\circ X=X\circ L_g$ for every $g\in G$. We can similarly define *right-invariant* vector fields, but most often we will be considering the left-invariant type. Since L_g is a diffeomorphism, the *push-out* of an arbitrary smooth vector field X on G can be defined as the vector field $L_{g*}X=dL_g\circ X\circ L_{g^{-1}}$. In this way, the condition of X to be left-invariant can be neatly expressed as $L_{g*}X=X$ for every $g\in G$.

Let $\mathfrak{g}$ denote the set of left invariant vector fields on G. It is clear that $\mathfrak{g}$ is a real vector space. Moreover, the map $X\in\mathfrak{g}\mapsto X_1$ defines a linear isomorphism between $\mathfrak{g}$ and the tangent space to

G at the identity T_1G, since any left invariant vector field is completely defined by its value at the identity. This implies that $\dim \mathfrak{g} = \dim G$. Every left invariant vector field X in G is smooth. This can be seen as follows. Let f be a smooth function defined on a neighborhood of 1 in G, and let $\gamma : (-\epsilon, \epsilon) \to G$ be a smooth curve with $\gamma(0) = 1$ and $\gamma'(0) = X_1$. Then the value of X on f is given by

$$X_g(f) = dL_g(X_1)(f) = X_1(f \circ L_g) = \frac{d}{dt}\Big|_{t=0} f(g\gamma(t)) = \frac{d}{dt}\Big|_{t=0} f \circ \mu(g, \gamma(t)),$$

and hence, it is a smooth function of g. Since the elements of $\mathfrak{g}$ are smooth vector fields, the bracket bewteen any two of them is defined. We end this discussion by observing that the bracket of X, $Y \in \mathfrak{g}$ is an element of $\mathfrak{g}$, for

$$L_{g*}[X, Y] = [L_{g*}X, L_{g*}Y] = [X, Y],$$

for every $g \in G$, due to Proposition 0.3.12.

The discussion in the previous paragraph shows that to any Lie group G is naturally associated a (real) Lie algebra $\mathfrak{g}$ consisting of the left invariant vector fields on G. This Lie algebra is the infinitesimal object associated to G and, as we shall see, completely determines its local structure.

0.4.3 Examples (a) $\mathbf{R}^n$ and T^n have the same Lie algebra, namely, the n-dimensinal Abelian Lie algebra.

(b) The Lie algebra of the direct product $G \times H$ is the direct sum of Lie algebras $\mathfrak{g} \oplus \mathfrak{h}$.

(c) G and G° have the same Lie algebra.

(d) The Lie algebra of a discrete group is $\{0\}$.

(e) The Lie algebra of $GL(n, \mathbf{R})$ is $\mathfrak{gl}(n, \mathbf{R})$ and that of $O(n)$ is $\mathfrak{so}(n)$. ⋆

The exponential map, subgroups and homomorphisms

Let G be a Lie group, and let $\mathfrak{g}$ denote its Lie algebra. Given $X \in \mathfrak{g}$, there exists an integral curve $\varphi_X : (-\epsilon, \epsilon) \to G$ of X with $\varphi(0) = 1$; namely, $\varphi'_X(t) = X_{\varphi_X(t)}$. Since

$$\frac{d}{dt}\Big|_{t=0} L_g(\varphi_X(t)) = d(L_g)_1(X_1) = X_g,$$

we have that $L_g \circ \varphi_X$ is the unique integral curve of X starting at g. In particular, by taking $g = \varphi(s)$ with s very close to ϵ, this shows that φ_X can be extended beyond ϵ. It follows that X is a complete vector field; namely, φ_X is defined on $\mathbf{R}$. Now $t \mapsto \varphi_X(s + t)$ for $s \in R$ is an integral curve of X with initial point $\varphi_X(s)$, and hence, by the uniqueness of integral curves,

$$\varphi_X(s + t) = \varphi_X(s)\varphi_X(t),$$

for every s, $t \in \mathbf{R}$. Because of this, we say that $\varphi_X : \mathbf{R} \to G$ is a *one-parameter subgroup* of G.

The *exponential map* of G is the map $\exp : \mathfrak{g} \to G$ defined by $\exp X = \varphi_X(1)$. We have $\varphi_{sX}(t) = \varphi_X(st)$, because $\frac{d}{dt}|_{t=0}\varphi_X(st) = s\varphi'_X(0) = sX$. This implies that $\varphi_X(t) = \varphi_{tX}(1) = \exp(tX)$, that is, every one-parameter subgroup factors through the exponential map.

The exponential map is smooth, as this follows from the smooth dependence of solutions of ordinary differential equations on initial conditions. Moreover, $d\exp_0 : T_0\mathfrak{g} \cong \mathfrak{g} \to T_1G \cong \mathfrak{g}$ is the identity, since

$$d\exp_0(X) = \frac{d}{dt}\Big|_{t=0} \exp(tX) = \varphi'_X(0) = X.$$

Thus, exp is a diffeomorphism from a neighborhood of 0 in $\mathfrak{g}$ onto a neighborhood of 1 in G.

0.4.4 Example The exponential map $\exp : \mathfrak{gl}(n,\mathbf{R}) \to GL(n,\mathbf{R})$ is the exponentiation of matrices:

$$\exp A = e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots$$

for all $A \in \mathfrak{gl}(n,\mathbf{R})$. In fact, for $\varphi_A(t) = e^{tA}$ we have that $\varphi_A'(t) = e^{tA}A = (dL_{\varphi_A(t)})A$ is the left-invariant vector field determined by A, so φ_A is its flow. Similarly for $\exp : \mathfrak{gl}(V) \to GL(V)$ where V is any real or complex vector space. ⋆

0.4.5 Remark In general, the exponential map is not a global diffeomorphism (take G compact), not a homomorphism (take G non-Abelian), not surjective (take $G = SL(2,\mathbf{R})$). We shall see on page 59 that exp is surjective if G is compact and connected.

The connected component of 1 in G, G°, is an open subgroup of G. G° is generated as a group by any neighborhood U of 1 (in fact, replace U by $U \cap U^{-1}$ in order to have $U = U^{-1}$; define $V = \cup_{n\geq 1} U^n$ and consider the equivalence relation $g \sim g'$ if and only if $g^{-1}g' \in V$; then the equivalence classes are open, whence, $V = G^\circ$). In particular, G° is generated by $\exp[\mathfrak{g}]$. This fact has major implications in the relation between $\mathfrak{g}$ and G.

Let G be a Lie group. A subgroup H of G is called a *Lie subgroup* of G if H is an (immersed) submanifold of G, and a Lie group with respect to the operations induced from G. If $\mathfrak{g}$ is a Lie algebra, a subspace $\mathfrak{h}$ of $\mathfrak{g}$ is called a *Lie subalgebra* if $\mathfrak{h}$ is closed under the bracket of $\mathfrak{g}$.

It is easy to see that if H is a Lie subgroup of G, then the Lie algebra of $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. Conversely, we have

0.4.6 Theorem (Lie) *Let G be a Lie group, and let $\mathfrak{g}$ denote its Lie algebra. If $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$, then there exists a unique connected Lie subgroup H of G such that the Lie algebra of H is $\mathfrak{h}$.*

Proof. This follows from the global version of Frobenius theorem. We have that $\mathfrak{h}$ is a subspace of T_1G. Let $\mathcal{D}$ be the left-invariant distribution on G defined by $\mathfrak{h}$. Then $\mathcal{D}$ is a smooth distribution, and the fact that $\mathfrak{h}$ is a subalgebra is equivalent to $\mathcal{D}$ being involutive. By Frobenius theorem, there is a unique maximal integral manifold of $\mathcal{D}$ passing through 1, which we call H. Then, for every $h \in H$, $h^{-1}H$ is also a maximal integral manifold of $\mathcal{D}$ passing through 1, which implies that $h^{-1}H = H$. It follows that H is a subgroup of G. Finally, the operations induced by G on H are smooth because H is an integral manifold of an involutive distribution (see Theorem 1.62 in [War83]). □

0.4.7 Remark A closed subgroup H of a Lie group G has a unique structure of Lie subgroup of G, and the underlying topology must be the induced topology, see [War83, p. 110].

A *(Lie group) homomorphism* between Lie groups G and H is map $\varphi : G \to H$ which is both a group homomorphism and a smooth map. φ is called a *isomorphism* if, in addition, it is a diffeomorphism. An *automorphism* of a Lie group is an isomorphism of the Lie group with itself. A *(Lie algebra) homomorphism* between Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ is a linear map $\Phi : \mathfrak{g} \to \mathfrak{h}$ which preserves brackets. Φ is called a *isomorphism* if, in addition, it is bijective. An *automorphism* of a Lie algebra is an isomorphism of the Lie algebra with itself.

A homomorphism $\varphi : G \to H$ between Lie groups induces a homomorphism $d\varphi : \mathfrak{g} \to \mathfrak{h}$ between the corresponding Lie algebras. Indeed, if X is a left invariant vector field on G, let Y be the unique left invariant vector field on H such that $Y_1 = d\varphi_1(X_1)$. Then

$$Y_{\varphi(g)} = d(L_{\varphi(g)})_1(Y_1) = d(L_{\varphi(g)} \circ \varphi)_1(X_1) = d(\varphi \circ L_g)_1(X_1) = d\varphi_g(X_g),$$

so that X and $Y|\varphi(G)$ are φ-related. Define $Y = d\varphi(X)$. Now, if $X' \in \mathfrak{g}$, then X' and $\varphi(X')$ are φ-related. Therefore $[X, X']$ and $[d\varphi(X), d\varphi(X')]|\varphi(G)$ are φ-related and thus

$$d\varphi([X, X']) = [d\varphi(X), d\varphi(X')].$$

This shows that $d\varphi$ is a Lie algebra homomorphism.

Let $\varphi : G \to H$ be a homomorphism between Lie groups. Then, for a left invariant vector field X on G, $t \mapsto \varphi(\exp^G(tX))$ is a one-parameter subgroup of H with $\frac{d}{dt}|_{t=0}\varphi(\exp^G tX) = d\varphi(X)$. It follows that

$$\varphi \circ \exp^G X = \exp^H \circ d\varphi(X), \tag{0.4.8}$$

for every X. In particular, if K is a Lie subgroup of G, then the inclusion map $i : K \to G$ is a Lie group homomorphism, so that the exponential map of G restricts to the exponential map of K, and the connected component of K is generated by $\exp^G[\mathfrak{k}]$, where $\mathfrak{k}$ is the Lie algebra of K. Since K is an integral manifold of an involutive distribution (compare Theorem0.4.6), it follows also that

$$\mathfrak{k} = \{X \in \mathfrak{g} : \exp^G(tX) \in K, \quad \text{for all } t \in \mathbf{R}\}.$$

0.4.9 Lemma *Let $\varphi : G \to H$ be a homomorphism between Lie groups. Consider the induced homomorphism between the corresponding Lie algebras $d\varphi : \mathfrak{g} \to \mathfrak{h}$. Then:*

a. $d\varphi$ is injective if and only if the kernel of φ is discrete.
b. $d\varphi$ is surjective if and only if $\varphi(G^\circ) = H^\circ$.
c. $d\varphi$ is bijective if and only if φ is a covering (here we assume G and H connected).

Proof. (a) $\ker\varphi$ is a closed normal subgroup of G, and its Lie algebra is $\ker d\varphi$.

(b) Since $\varphi \circ \exp = \exp \circ d\varphi$, and G° is generated by $\exp[\mathfrak{g}]$, $\varphi(G^\circ)$ is the subgroup of H° generated by $\exp[d\varphi(\mathfrak{g})]$.

(c) Suppose G, H connected, $d\varphi : \mathfrak{g} \to \mathfrak{h}$ an isomorphism. Then φ is surjective by (b). Let U be a neighborhood of 1 in G such that $\varphi : U \to \varphi(U) := V$ is a diffeomorphism. We can choose U so that $U \cap \ker d\varphi = \{1\}$ by (a). Then $\varphi^{-1}(V) = \cup_{n\in\ker\varphi} nU$ (disjoint union), and, since $\varphi \circ L_n = \varphi$ for $n \in \ker\varphi$, we also have that $\varphi|nU$ is a diffeomorphism onto V. This shows that φ is a covering. The other half of the statement is clear. □

0.4.10 Theorem *Let G_1, G_2 be Lie groups, and assume that G_1 is connected and simply-connected. Then, given a homomorphism $\Phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ between the Lie algebras, there exists a unique homomorphism $\varphi : G_1 \to G_2$ such that $d\varphi = \Phi$.*

Proof. The graph of Φ, $\mathfrak{h} = \{(X, \Phi(X)) : X \in \mathfrak{g}_1$ is a subalgebra of $\mathfrak{g}_1 \oplus \mathfrak{g}_2$. Let H be the subgroup of $G_1 \times G_2$ defined by $\mathfrak{h}$ (Theorem 0.4.6). Consider the projections

$$\Phi_i : \mathfrak{g}_1 \oplus \mathfrak{g}_2 \to \mathfrak{g}_i, \qquad \varphi_i : G_1 \times G_2 \to G_i,$$

for $i = 1, 2$. Since $\Phi_1|\mathfrak{h} : \mathfrak{h} \to \mathfrak{g}_q$ is an isomorphism, we have that $\Phi = \Phi_2 \circ (\Phi_1|\mathfrak{h})^{-1}$ and $\varphi_1 : H \to G_1$ is a covering. Since G_1 is simply-connected, $\varphi_1|H : H \to G_1$ is an isomorphism of Lie groups, and we can thus define $\varphi = \varphi_2 \circ (\varphi_1)^{-1}$. This proves the existence part. The uniqueness part comes from the fact that $d\varphi = \Phi$ specifies φ in a neighborhood of 1 (by using the exponential map), and G_1 is generated by this neighborhood. □

The adjoint representation

Let G be a Lie group, and denote its Lie algebra by $\mathfrak{g}$. The noncommutativity of G is organized by the adjoint representation. In order to introduce it, let $g \in G$, and define a map $\mathrm{Inn}_g : G \to G$ by $\mathrm{Inn}_g(x) = gxg^{-1}$. Then Inn_g is an automorphism of G, which is called the *inner automorphism defined by* g. The differential $d(\mathrm{Inn}_g) : \mathfrak{g} \to \mathfrak{g}$ defines an automorphism of $\mathfrak{g}$, which we denote by Ad_g. Then

$$\mathrm{Ad}_g X = \frac{d}{dt}\Big|_{t=0} \mathrm{Inn}(g)(\exp tX) = \frac{d}{dt}\Big|_{t=0} g \exp tX g^{-1}.$$

0.4.11 Example In case $G = GL(n, \mathbf{R})$ we have (cf. example 0.4.4)

$$\begin{aligned} \mathrm{Ad}_g X &= \frac{d}{dt}\Big|_{t=0} g e^{tX} g^{-1} \\ &= \frac{d}{dt}\Big|_{t=0} e^{t(gXg^{-1})} \\ &= gXg^{-1}. \end{aligned}$$

⋆

Now we have a homomorphism

$$\mathrm{Ad} : g \in G \to \mathrm{Ad}_g \in \mathbf{GL}(\mathfrak{g}),$$

which is called the *adjoint representation* of G on $\mathfrak{g}$. We have

$$\begin{aligned} \mathrm{Ad}_g X &= (dL_g)_1 (dR_{g^{-1}})_1 X_1 \\ &= (dR_{g^{-1}})_1 (dL_g)_1 X_1 \\ &= (dR_{g^{-1}})_1 (X_g) \\ &= (dR_g^{-1} \circ X \circ R_g)_1 \\ &= \big((R_{g^{-1}})_* X\big)_1 \end{aligned}$$

Finally, the differential $d(\mathrm{Ad})$ defines the *adjoint representation* of $\mathfrak{g}$ on $\mathfrak{g}$:

$$\mathrm{ad} : X \in \mathfrak{g} \to \mathrm{ad}_X = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp tX} \in \mathfrak{gl}(\mathfrak{g}).$$

Since $\varphi_t = R_{\exp tX}$ is the flow of X, we get

$$\mathrm{ad}_X Y = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp tX} Y = \frac{d}{dt}\Big|_{t=0} \big((R_{\exp(-tX)})_* Y\big)_1 = (L_X Y)_1 = [X, Y].$$

As an important special case of (0.4.8) we have (recall example 0.4.4)

$$\mathrm{Ad}_{\exp X} = e^{\mathrm{ad}_X}$$

for all $X \in \mathfrak{g}$.

0.4.12 Lemma $[X, Y] = 0$ *if and only if* $\exp X \exp Y = \exp Y \exp X$ *for all* X, $Y \in \mathfrak{g}$. *In that case,* $\exp(t(X+Y)) = \exp tX \exp tY$ *for all* $t \in \mathbf{R}$. *It follows that a connected Lie group is Abelian if and only if its Lie algebra is Abelian.*

Proof. The first assertion is a special case of Proposition 0.3.13 using that $\varphi_t = R_{\exp tX}$ is the flow of X and $\psi_s = R_{\exp sY}$ is the flow of Y. The second one follows from noting that both $t \mapsto \exp(t(X+Y))$ and $t \mapsto \exp tX \exp tY$ are one-parameter groups with initial speed $X+Y$. Finally, we have seen that $\mathfrak{g}$ is Abelian if and only if $\exp[\mathfrak{g}]$ is Abelian, but the latter generates G°. □

0.4.13 Theorem *Every connected Abelian Lie group G is isomorphic to $\mathbf{R}^{n-k} \times T^k$. In particular, a simply-connected connected Abelian Lie group is isomorphic to $\mathbf{R}^n$ and a compact connected Abelian Lie group is isomorphic to T^n.*

Proof. It follows from Lemma 0.4.12 that $\mathfrak{g}$ is Abelian and $\exp : \mathfrak{g} \to G$ is a homomorphism, where $\mathfrak{g} \cong \mathbf{R}^n$ as a Lie group, thus exp is a smooth covering by Lemma 0.4.9(c). Hence G is isomorphic to $\mathbf{R}^n$ quotiented by the discrete group $\ker \exp$. □

Lie transformation groups

As mentioned above, Lie groups serve to model transformations of manifolds. Let G be a Lie group and let M be a smooth manifold. A *smooth action* of G on M, also called a *Lie transformation group*, is a homomorphism Φ of G into the group of diffeomorphisms of M such that the map

$$G \times M \to M, \qquad (g,p) \mapsto \Phi(g)p$$

is smooth. We usually write gp for $\Phi(g)p$. In this case one says that G acts on M by diffeomorphisms. The *isotropy group* at $p \in M$ is the subgroup G_p of G consisting of all elements that fix p, namely, $G_p = \{\, g \in G \mid gp = p \,\}$. The *orbit* through $p \in M$ is the subset Gp of points of M that can be attained from p under the action of G, namely, $Gp = \{\, gp \mid g \in G \,\}$. Note that the orbits of an action partition the space into equivalence classes. The quotient space is also called *orbit space*.

0.4.14 Lemma *Let $\sim$ be an equivalence relation on a topological space X such that the natural projection $\pi : X \to X/\sim$ mapping each $x \in X$ to its equivalence class $[x]$ is an open map. Then the quotient space $X/\sim$ is Hausdorff if and only if $\sim$ is closed in $X \times X$.*

Proof. Note that $[x] \neq [y]$ if and only if $(x,y) \notin \sim$. Also, $\sim$ is closed if and only if for such (x,y) there is an open neighborhood in $X \times X$, which can be assumed of the form $V \times W$ for V, W open neighborhoods of x, y in X, resp., which does not meet $\sim$. However, the existence of such neighborhoods V, W is the same as separating $[x]$, $[y]$ by open sets since π is continuous and open. □

An action of G on M is called *proper* if the induced map

$$G \times M \to M \times M, \qquad (g,p) \mapsto (gp,p) \tag{0.4.15}$$

is a proper map (compare page 6). It is equivalent to require that for all compact subsets K, $L \subset M$, the set $\{\, g \in G \mid gK \cap L \neq \varnothing \,\}$ be compact. In this form, one easily sees that this definition extends the one given previously for discrete groups (see page 9). Note that properness of the action is automatic if G is a compact Lie group.

0.4.16 Theorem *If M is a smooth manifold and G is a Lie group acting freely and properly on M, then the quotient space $\bar{M} = G\backslash M$ endowed with the quotient topology admits a natutal structure of smooth manifold such that the projection $\pi : M \to \bar{M}$ is a (surjective) submersion. Moreover* $\dim \bar{M} = \dim M - \dim G$.

Proof. We start by noting that π is an open map, as for an open set V of M we have $\pi^{-1}(\pi(V)) = \cup_{g\in G} gV$ is a union of open sets and thus open. It follows that a projection of a contable basis of open sets of M yields a countable basis of open sets of $\bar{M}$. Moreover, $\bar{M}$ is Hausdorff since the range of the proper map (0.4.15) is closed and thus we can apply Lemma 0.4.14.

Fix $p \in M$. The map $\omega_p : G \to M$, $\omega_p(g) = gp$ is smooth by definition of an action, injective by freeness of the action and proper by properness of the action. It is also an immersion, as we show now. Since

$$\omega_p \circ L_g = \Phi(g) \circ \omega_p$$

and $L_g : G \to G$, $\Phi(g) : M \to M$ are diffeomorphisms, it suffices to check that ω_p is an immersion at $1 \in G$. Let $X \in \mathfrak{g} \cong T_1G$. Then

$$X_p^* := d\omega_p(X) = \frac{d}{dt}\Big|_{t=0} (\exp tX)p$$

defines a smooth vector field on M whose flow is $\tilde{\varphi}_t = \Phi(\exp tX)$, so $X_p^* = 0$ if and only if the integral curve through p is constant, namely, $\Phi(\exp tX)p = p$ for all $t \in \mathbf{R}$ which, due to freeness, says that $X = 0$. Now ω_p is a proper injective immersion and hence its image, the orbit Gp, is a properly embedded submanifold of M.

Let us construct a local chart of $\bar{M}$ around $\bar{p} = \pi(p) = Gp \in \bar{M}$. There is a local chart (U, φ) of M adapted to Gp around p. Suppose $\dim M = n + k$, $\dim Gp = n$. We may assume that $\varphi(p) = 0$, $\varphi(U) \subset \mathbf{R}^{n+k} = \mathbf{R}^n \times \mathbf{R}^k$ is a product neighborhood $V \times W$ of 0, where where $V = \varphi(U) \cap \mathbf{R}^n$ and W is a neighborhood of 0 in $\mathbf{R}^k$. Define a smooth map $F : G \times W \to M$ by $F(g, y) = g\varphi^{-1}(y)$. Then $dF_{(1,0)}$ maps T_1G onto $T_p(Gp)$, which equals $d(\varphi^{-1})_0(\mathbf{R}^n)$, and it maps $T_0W = \mathbf{R}^k$ onto $d(\varphi^{-1})_0(\mathbf{R}^k)$. Since $d(\varphi^{-1})_0(\mathbf{R}^n) + d(\varphi^{-1})(\mathbf{R}^k)_0 = T_pM$, F is a local diffeomorphism at $(1, 0)$. By shrinking W and using that $F(g, y) = \Phi(g)F(1, y)$, we can ensure that F is a local diffeomorphism at every point of $G\times W$. Next we claim it is possible to further shrink W to arrange that F is injective and thus a diffeomorphism onto its image. Otherwise, there would be sequences (g_i), (h_i) in G, (y_i), (z_i) in W such that $y_i \to 0$, $z_i \to 0$, $g_i\varphi^{-1}(y_i) = h_i\varphi^{-1}(z_i)$ but $(g_i, y_i) \neq (h_i, z_i)$ for all i. Put $k_i := h_i^{-1}g_i \in G$. Since $(k_i\varphi^{-1}(y_i), \varphi^{-1}(y_i)) = (\varphi^{-1}(z_i), \varphi^{-1}(y_i)) \to (p, p)$, the $(k_i\varphi^{-1}(y_i), \varphi^{-1}(y_i))$ are eventually contained in a compact subset of $M \times M$ and thus, by properness of the action, the k_i are eventually contained in a compact subset of G; by passing to a subsequence, we may assume that $k_i \to k \in G$. Now $p = \lim k_i\varphi^{-1}(y_i) = kp$ which implies $k = 1$ by freeness of the action. However, this contradicts the local injectivity of F at $(1, 0)$, proving the claim. Now for $U = F(W)$ we have a diffeomorphism $\psi = F^{-1} : U \to G \times W$. Let $\psi_1 : U \to G$, $\psi_2 : U \to W$ denote the components of ψ. Note that U is a “fibered” neighborhood of Gp in the sense that the nearby orbits Gq map to fibers of the form $G \times \{y\}$ where $y = \psi_2(q) \in W$. Note also that $S := \psi^{-1}(\{1\} \times W)$ is a “slice” near Gp in the sense that S meets each orbit in W in exactly one point. The map ψ is G-equivariant in the sense that $\psi(gq) = (g\psi_1(q), \psi_2(q))$ for $g \in G$, $q \in U$. Now ψ_2 induces a homeomorphism $\bar{\psi}_2 : \pi(U) \to W$ from the open neighborhood $\pi(U)$ of $\bar{p}$ in $\bar{M}$ onto the open neighborhood W of 0 in $\mathbf{R}^k$, which we take as a local chart of $\bar{M}$.

We can cover $\bar{M}$ with local charts of this form. Suppose $\bar{\psi}'_2 : \pi(U') \to W'$ is another chart coming from $\psi' = (\psi'_1, \psi'_2) : U' \to G \times W'$ such that $\pi(U) \cap \pi(U') \neq \varnothing$. Let $y \in \bar{\psi}_2(\pi(U) \cap \pi(U'))$. Then $\bar{\psi}_2^{-1}(y) = \pi(\psi^{-1}(1, y))$, so the transition map

$$\bar{\psi}'_2\bar{\psi}_2^{-1}(y) = (\bar{\psi}'_2\pi)\psi^{-1}\iota(y) = \psi'_2\psi^{-1}\iota(y)$$

is smooth, where $\iota(y) = (1, y)$. This proves that we have a smooth atlas.

The commutative diagram

$$\begin{array}{ccc} M \supset U & \xrightarrow{\psi} & G \times W \\ \pi \big\downarrow & & \big\downarrow \\ \bar{M} \supset \pi(U) & \xrightarrow[\bar{\psi}_2]{} & W \end{array}$$

shows that π is a submersion. □

0.4.17 Remarks (a) In the notation of the preceding theorem, the smooth map $s : \pi(U) \to M$ defined by $s(\bar{q}) = \psi^{-1}(1, \bar{\psi}_2(\bar{q}))$ has image S and satisfies $\pi \circ s = \mathrm{id}_{\pi(U)}$. A smooth map $s : \mathcal{O} \to M$, where $\mathcal{O}$ is an open set of $\bar{M}$, satisfying $\pi \circ s = \mathrm{id}_{\mathcal{O}}$ is called a (smooth) *local section* of $\pi : M \to \bar{M}$.

(b) The proof of the preceding theorem has indeed revealed more, namely, $\pi : M \to \bar{M}$ is a principal G-bundle. A smooth map $\pi : M \to B$ between smooth manifolds is called a *principal G-bundle*, where G is a Lie group, if M is equipped with a free, *right* action of G and B can be covered by neighborhoods $\mathcal{O}$ such that $\pi^{-1}(\mathcal{O})$ is diffeomorphic to $\mathcal{O} \times G$, where fibers of π are mapped to fibers of $\mathcal{O} \times G \to \mathcal{O}$, and the action of G on M corresponds to its action by right multiplication on the second factor of $\mathcal{O} \times G$. The smooth structure constructed on $\bar{M}$ in the theorem is the unique one that makes $\pi : M \to \bar{M}$ into a smooth principal G-bundle.

(c) A map $\bar{f} : \bar{M} \to N$ is smooth if and only if $f := \pi \circ \bar{f} : M \to N$ is smooth. This essentially follows from the commutative diagram in the proof.

As the most important application of Theorem 0.4.16, let G be a Lie group and let H be a closed subgroup. Then H acts on G by right multiplication as follows:

$$\Phi : H \times G \to G, \qquad \Phi(h)g := R_{h^{-1}}g$$

(note that the inverse in h^{-1} is necessary to have an action "on the left", as we have defined). This action is clearly free. It is also proper, because given compact subsets K, $L \subset G$, the set $\{ h \in H \mid hK \cap L \neq \varnothing \}$ coincides with $L^{-1}K \cap H$, which is compact. The orbits of this action coincide with the co-classes of G module H, namely, gH for $g \in G$. Hence

0.4.18 Theorem *If G is a Lie group and H is a closed subgroup of G, then there exists a natural structure of smooth manifold on the quotient G/H such that the projection $G \to G/H$ is a submersion. Moreover,* $\dim G/H = \dim G - \dim H$.

Let G be a Lie group acting by diffeomorphisms on a smooth manifold M. We say that the action is *transitive* if for any p, $q \in M$ there exists $g \in G$ such that $gp = q$; equivalently, there is only one orbit of G in M. In this case, we say that M is *homogeneous* under G or that M is a *homogeneous space*. It is clear that G/H as in Theorem 0.4.18 is always homogeneous under G, where G acts by left multiplication: given g_1H, $g_2H \in G/H$, the element $g_2g_1^{-1}$ maps one point to the other. Conversely:

0.4.19 Theorem *Let G act transitively on M. Then, for any $p \in M$, the orbit map $\omega_p : G \to M$, $\omega_p(g) = gp$ induces a diffeomorphism $G/G_p \to M$.*

Proof. Since G_p is a closed subgroup of G, it is a Lie subgroup of G (Remark 0.4.7) and thus G/G_p is a smooth manifold. As is easy to see, the map $\bar{\omega}_p : gG_p \in G/G_p \mapsto gp \in M$ is well defined, bijective and smooth. As in the proof of Theorem 0.4.16, one shows that $\bar{\omega}_p$ is an immersion at

$1G_p$ and thus an immersion everywhere by equivariance. This already implies $\dim G/G_p \leq \dim M$, and the image of $\bar{\omega}_p$ is a submanifold of M, but the strictly inequality cannot hold as $\bar{\omega}_p$ is bijective and the image of a smooth map from a smooth manifold into a strictly higher dimensional smooth manifold has null measure (this result follows from the statement that the image of a smooth map $\mathbf{R}^n \to \mathbf{R}^{n+k}$ with $k > 0$ has null measure *and the second-countability of smooth manifolds*). It follows that $\bar{\omega}_p$ is a local diffeomorphism and hence a diffeomorphism. □

0.4.20 Corollary *The smooth structure in G/H constructed in Theorem 0.4.18 is the unique one that makes the action of G on G/H by left multiplication smooth.*

0.5 Vector bundles ★

CHAPTER 1

Riemannian manifolds

1.1 Introduction

A Riemannian metric is a family of smoothly varying inner products on the tangent spaces of a smooth manifold. Riemannian metrics are thus infinitesimal objects, but they can be used to measure distances on the manifold. They were introduced by Riemann in his seminal work [Rie53] in 1854. At that time, the concept of a manifold was extremely vague and, except for some known global examples, most of the work of the geometers focused on local considerations, so the modern concept of a Riemannian manifold took quite some time to evolve to its present form. We point out the seemingly obvious fact that a given smooth manifold can be equipped with many different Riemannian metrics. This is really one of the great insights of Riemann, namely, the separation between the concepts of space and metric.

This chapter is mainly concerned with examples.

1.2 Riemannian metrics

Let M be a smooth manifold. A *Riemannian metric* g on M is a smoothly varying family of inner products on the tangent spaces of M. Namely, g associates to each $p \in M$ a positive definite symmetric bilinear form on T_pM,

$$g_p : T_pM \times T_pM \to \mathbf{R},$$

and the smoothness condition on g refers to the fact that the function

$$p \in M \mapsto g_p(X_p, Y_p) \in \mathbf{R}$$

must be smooth for every locally defined smooth vector fields X, Y in M. A *Riemannian manifold* is a pair (M, g) where M is a differentiable manifold and g is a Riemannian metric on M. Later on (but not in this chapter), we will often simplify the notation and refer to M as a Riemannian manifold where the Riemannian metric is implicit.

Let (M, g) be a Riemannian manifold. If $(U, \varphi = (x^1, \ldots, x^n))$ is a chart of M, a local expression for g can be given as follows. Let $\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\}$ be the coordinate vector fields, and let $\{dx^1, \ldots, dx^n\}$ be the dual 1-forms. For $p \in U$ and u, $v \in T_pM$, we write

$$u = \sum_i u^i \frac{\partial}{\partial x^i}\Big|_p \quad \text{and} \quad v = \sum_j v^j \frac{\partial}{\partial x^j}\Big|_p.$$

Then, by bilinearity,

$$\begin{aligned} g_p(u,v) &= \sum_{i,j} u^i v^j g_p\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \\ &= \sum_{i,j} g_{ij}(p) u^i v^j, \end{aligned}$$

where we have set

$$g_{ij}(p) = g_p\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right).$$

Note that $g_{ij} = g_{ji}$. Hence we can write

$$g = \sum_{i,j} g_{ij}\, dx^i \otimes dx^j = \sum_{i\leq j} \tilde{g}_{ij}\, dx^i dx^j, \tag{1.2.1}$$

where $\tilde{g}_{ii} = g_{ii}$, $\tilde{g}_{ij} = 2g_{ij}$ if $i < j$, and $dx^i dx^j = \frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i)$.

Next, let $(U', \varphi' = (x'^1, \ldots, x'^n))$ be another chart of M such that $U \cap U' \neq \varnothing$. Then

$$\frac{\partial}{\partial x'^i} = \sum_k \frac{\partial x^k}{\partial x'^i} \frac{\partial}{\partial x^k},$$

so the relation between the local expressions of g with respect to (U, φ) and (U', φ') is given by

$$g'_{ij} = g\left(\frac{\partial}{\partial x'^i}, \frac{\partial}{\partial x'^j}\right) = \sum_{k,l} \frac{\partial x^k}{\partial x'^i} \frac{\partial x^l}{\partial x'^j}\, g_{kl}.$$

1.2.2 Examples (a) The canonical Euclidean metric is expressed in Cartesian coordinates by $g = dx^2 + dy^2$. Changing to polar coordinates $x = r\cos\theta$, $y = r\sin\theta$ yields that

$$dx = \cos\theta dr - r\sin\theta d\theta \quad \text{and} \quad dy = \sin\theta dr + r\cos\theta d\theta,$$

so

$$\begin{aligned} g &= dx^2 + dy^2 \\ &= (\cos^2\theta dr^2 + r^2\sin^2\theta d\theta^2 - 2r\sin\theta\cos\theta dr d\theta) \\ &\qquad +(\sin^2\theta dr^2 + r^2\cos^2\theta d\theta^2 + 2r\sin\theta\cos\theta dr d\theta) \\ &= dr^2 + r^2 d\theta^2. \end{aligned}$$

(b) A classical example is the surface of revolution parametrized by

$$\mathbf{x}(r,\theta) = (a(r)\cos\theta, a(r)\sin\theta, b(r)),$$

where $a > 0$, b are smooth functions defined on some interval and the generatrix $\gamma(r) = (a(r), 0, b(r))$ has $||\gamma'||^2 = (a')^2 + (b')^2 = 1$, equipped with the metric g induced from $\mathbf{R}^3$. Namely, the tangent spaces to the surface are subspaces of $\mathbf{R}^3$, so we can endow them with inner products just by taking the restrictions of the Euclidean dot product in $\mathbf{R}^3$. The tangent spaces are spanned by the partial derivatives $\mathbf{x}_r = (\frac{\partial x}{\partial r}, \frac{\partial y}{\partial r}, \frac{\partial z}{\partial r})$, $\mathbf{x}_\theta = (\frac{\partial x}{\partial \theta}, \frac{\partial y}{\partial \theta}, \frac{\partial z}{\partial \theta})$, and then $g = (\mathbf{x}_r \cdot \mathbf{x}_r)\, dr^2 + 2(\mathbf{x}_r \cdot \mathbf{x}_\theta)\, dr d\theta + (\mathbf{x}_\theta \cdot \mathbf{x}_\theta)\, d\theta^2$. Equivalently, from

$$\begin{aligned} dx &= a'(r)\cos\theta\, dr - a(r)\sin\theta\, d\theta \\ dy &= a'(r)\sin\theta\, dr + a(r)\cos\theta\, d\theta \\ dz &= b'(r)\, dr \end{aligned}$$

we obtain

$$\begin{aligned} g &= dx^2 + dy^2 + dz^2 \\ &= dr^2 + a(r)^2\, d\theta^2. \end{aligned}$$

⋆

The functions g_{ij} are smooth on U and, for each $p \in U$, the matrix $(g_{ij}(p))$ is symmetric and positive-definite. Conversely, a Riemannian metric in U can be obviously specified by these data.

1.2.3 Proposition *Every smooth manifold can be endowed with a Riemannian metric.*

Proof. Let $M = \cup_\alpha U_\alpha$ be a covering of M by domains of charts $\{(U_\alpha, \varphi_\alpha)\}$. For each α, consider the Riemannian metric g_α in U_α whose local expression $((g_\alpha)_{ij})$ is the identity matrix. Let $\{\rho_\alpha\}$ be a smooth partition of unity of M subordinate to the covering $\{U_\alpha\}$, and define

$$g = \sum_\alpha \rho_\alpha g_\alpha.$$

Since the family of supports of the ρ_α is locally finite, the above sum is locally finite, and hence g is well defined and smooth, and it is bilinear and symmetric at each point. Since $\rho_\alpha \geq 0$ for all α and $\sum_\alpha \rho_\alpha = 1$, it also follows that g is positive definite, and thus is a Riemannian metric in M. □

The proof of the preceding proposition suggests the fact that there exists a vast array of Riemannian metrics on a given smooth manifold. Even taking into account equivalence classes of Riemannian manifolds, the fact is that there many uninteresting examples of Riemannian manifolds, so an important part of the work of the differential geometer is to sort out relevant families of examples.

Let (M, g) and (M', g') be Riemannian manifolds. A *isometry* between (M, g) and (M', g') is diffeomorphism $f : M \to M'$ whose differential is a linear isometry between the corresponding tangent spaces, namely,

$$g_p(u, v) = g'_{f(p)}(df_p(u), df_p(v)),$$

for every $p \in M$ and u, $v \in T_pM$. We say that (M, g) and (M', g') are *isometric Riemannian manifolds* if there exists an isometry between them. This completes the definition of the category of Riemannian manifolds and isometric maps. Note that the set of all isometries of a Riemannian manifold (M, g) forms a group, called the *isometry group* of (M, g), with respect to the operation of composition of mappings, which we will denote by $\mathrm{Isom}(M, g)$. Here we quote without proof the following important theorem [MS39].

1.2.4 Theorem (Myers-Steenrod) *The isometry group* $\mathrm{Isom}(M, g)$ *of a Riemannian manifold* (M, g) *has the structure of a Lie group with respect to the compact-open topology. Its isotropy subgroup at an arbitrary fixed point is compact. Moreover,* $\mathrm{Isom}(M, g)$ *is compact if* M *is compact.*

The isometry group is a Riemannian-geometric invariant in the sense that if $f : (M, g) \to (M', g)$ is an isometry between Riemannian manifolds, then $\alpha \mapsto f \circ \alpha \circ f^{-1}$ defines an isomorphism $\mathrm{Isom}(M, g) \to \mathrm{Isom}(M', g')$.

A *local isometry* from (M, g) into (M', g') is a smooth map $f : M \to M'$ satisfying the condition that every point $p \in M$ admits a neighborhood U such that the restriction of f to U is an isometry onto its image. In particular, f is a local diffeomorphism. Note that a local isometry which is bijective is an isometry.

1.3 Examples

The Euclidean space

The Euclidean space is $\mathbf{R}^n$ equipped with its standard scalar product. The essential feature of $\mathbf{R}^n$ as a smooth manifold is that, since it is *the* model space for finite dimensional smooth manifolds, it admits a global chart given by the identity map. Of course, the identity map establishes canonical isomorphisms of the tangent spaces of $\mathbf{R}^n$ at each of its points with $\mathbf{R}^n$ itself. Therefore an arbitrary Riemannian metric in $\mathbf{R}^n$ can be viewed as a smooth family of inner products in $\mathbf{R}^n$. In particular, by taking the constant family given by the standard scalar product, we get the canonical Riemannian structure in $\mathbf{R}^n$. In this book, unless explicitly stated, we will always use its canonical metric when referring to $\mathbf{R}^n$ as a Riemannian manifold.

If $(x_1, \ldots, x_n)$ denote the standard coordinates on $\mathbf{R}^n$, then it is readily seen that the local expression of the canonical metric is

$$dx_1^2 + \cdots + dx_n^2. \tag{1.3.1}$$

More generally, if a Riemannian manifold (M, g) admits local coordinates such that the local expression of g is as in (1.3.1), then (M, g) is called *flat* and g is called a *flat metric* on M. Note that, if g is a flat metric on M, then the coordinates used to express g as in (1.3.1) immediately define a local isometry between (M, g) and Euclidean space $\mathbf{R}^n$.

Riemannian submanifolds and isometric immersions

Let (M, g) be a Riemannian manifold and consider an immersed submanifold $\iota : N \to M$. This means that N is a smooth manifold and ι is an injective immersion. Then the Riemannian metric g induces a Riemannian metric g_N in N as follows. Let $p \in N$. The tangent space T_pN can be viewed as a subspace of T_pM via the injective map $d\iota_p : T_pN \to T_{\iota(p)}M$. We define $(g_N)_p$ to be simply the restriction of g to this subspace, namely,

$$(g_N)_p(u, v) = g_{\iota(p)}(d\iota_p(u), d\iota_p(v)),$$

where u, $v \in T_pN$. It is clear that g_N is a Riemannian metric. We call g_N the *induced Riemannian metric in* N, and we call (N, g_N) a *Riemannian submanifold* of (M, g).

Note that the definition of g_N makes sense even if ι is an immersion that is not necessarily injective. In general, we call g_N the *pulled-back metric*, write $g_N = \iota^*g$, and say that $\iota : (N, g_N) \to (M, g)$ is an *isometric immersion* (of course, any immersion must be locally injective). On another note, an isometry $f : (M, g) \to (M', g')$ is a diffeomorphism satisfying $f^*(g') = g$.

A very important particular case is that of Riemannian submanifolds of Euclidean space (compare example 1.2.2(b)) Historically speaking, the study of Riemannian manifolds was preceded by the theory of curves and surfaces in $\mathbf{R}^3$. In the classical theory, one uses parametrizations instead of local charts, and these objects are called *parametrized curves* and *parametrized surfaces* since they usually already come with the parametrization. In the most general case, the parametrization is only assumed to be smooth. One talks about a *regular curve* or a *regular surface* if one wants the parametrization to be an immersion. Of course, in this case it follows that the parametrization is locally an embedding. This is good enough for the classical theory, since it is really concerned with local computations.

The sphere S^n

The canonical Riemannian metric in the sphere S^n is the Riemannian metric induced by its embedding in $\mathbf{R}^{n+1}$ as the sphere of unit radius. When one refers to S^n as a Riemannian manifold with its canonical Riemannian metric, sometimes one speaks of "the unit sphere", or "the metric sphere", or the "Euclidean sphere", or "the round sphere". One also uses the notation $S^n(R)$ to specify a sphere of radius R embedded in $\mathbf{R}^{n+1}$ with the induced metric. In this book, unless explicitly stated, we will always use the canonical metric when referring to S^n as a Riemannian manifold.

Product Riemannian manifolds

Let (M_i, g_i), where $i = 1, 2$, denote two Riemannian manifolds. Then the product smooth manifold $M = M_1 \times M_2$ admits a canonical Riemannian metric g, called the *product Riemannian metric*, given as follows. The tangent space of M at a point $p = (p_1, p_2) \in M_1 \times M_2$ splits as $T_pM = T_{p_1}M_1 \oplus T_{p_2}M_2$. Given u, $v \in T_pM$, write accordingly $u = u_1 + u_2$ and $v = v_1 + v_2$, and define

$$g_p(u, v) = g_{p_1}(u_1, v_1) + g_{p_2}(u_2, v_2).$$

It is clear that g is a Riemannian metric. Note that it follows from this definition that $T_{p_1}M_1 \oplus \{0\}$ is orthogonal to $\{0\} \oplus T_{p_2}M_2$. We will sometimes write that $(M, g) = (M_1, g_1) \times (M_2, g_2)$, or that $g = g_1 + g_2$.

It is immediate to see that Euclidean space $\mathbf{R}^n$ is the Riemannian product of n copies of $\mathbf{R}$.

Conformal Riemannian metrics

Let (M, g) be a Riemannian manifold. If f is a nowhere zero smooth function on M, then f^2g defined by

$$(f^2g)_p(u, v) = f^2(p)g_p(u, v),$$

where $p \in M$, u, $v \in T_pM$, is a new Riemannian metric on M which is said to be *conformal* to g. The idea behind this definition is that g and f^2g define the same angles between pairs of tangent vectors. We say that (M, g) is *conformally flat* if M can be covered by open sets on each of which g is conformal to a flat metric.

A particular case happens if f is a nonzero constant in which f^2g is said to be *homothetic* to g.

The real hyperbolic space $\mathbf{R}H^n$

To begin with, consider the Lorentzian inner product in $\mathbf{R}^{n+1}$ given by

$$\langle x, y \rangle = -x_0y_0 + x_1y_1 + \cdots + x_ny_n,$$

where $x = (x_0, \ldots, x_n)$, $y = (y_0, \ldots, y_n) \in \mathbf{R}^{n+1}$. We will write $\mathbf{R}^{1,n}$ to denote $\mathbf{R}^{n+1}$ with such a Lorentzian inner product. Note that if $p \in \mathbf{R}^{1,n}$ is such that $\langle p, p \rangle < 0$, then the restriction of $\langle , \rangle$ to $\langle p \rangle^\perp$ (the the orthogonal complement to p with regard to $\langle , \rangle$) is positive-definite (compare Exercise 15). Note also that the equation $\langle x, x \rangle = -1$ defines a two-sheeted hyperboloid in $\mathbf{R}^{1,n}$.

Now we can define the *real hyperbolic space* as the following submanifold of $\mathbf{R}^{1,n}$,

$$\mathbf{R}H^n = \{ x \in \mathbf{R}^{1,n} \mid \langle x, x \rangle = -1 \quad \text{and} \quad x_0 > 0 \},$$

equipped with a Riemannian metric g given by the restriction of $\langle , \rangle$ to the tangent spaces at its points. Since the tangent space of the hyperboloid at a point p is given by $\langle p \rangle^\perp$, the Riemannian metric g turns out to be well defined. Actually, this submanifold is sometimes called the *hyperboloid model of* $\mathbf{R}H^n$ (compare Exercises 3 and 4). This model brings about the duality between S^n and $\mathbf{R}H^n$ in the sense that one can think of the hyperboloid as the sphere of unit imaginary radius in $R^{1,n}$. Of course, as a smooth manifold, $\mathbf{R}H^n$ is diffeomorphic to $\mathbf{R}^n$.

Flat tori

A *lattice* Γ in $\mathbf{R}^n$ (or, more generally, in a real vector space) is the additive subgroup of $\mathbf{R}^n$ consisting of integral linear combinations of the vectors in a fixed basis. Namely, if $\{v_1, \ldots, v_n\}$ is a basis of $\mathbf{R}^n$, then it defines the lattice $\Gamma = \{ \sum_{j=1}^n m_j v_j \mid m_1, \ldots, m_n \in \mathbf{Z} \}$. For a given lattice Γ we consider the quotient group $\mathbf{R}^n/\Gamma$ in which two elements p, $q \in \mathbf{R}^n$ are identified if $q - p \in \Gamma$. We will show that $M = \mathbf{R}^n/\Gamma$ has the structure of a compact smooth manifold of dimension n diffeomorphic to a product of n copies of S^1, which we denote by T^n. Moreover there is a naturally defined flat metric g_Γ on M; the resulting Riemannian manifold is called a *flat torus*. We also denote it by (T^n, g_Γ).

Relevant for the topology of M will be the discreteness of Γ as an additive subgroup of $\mathbf{R}^n$, namely: any bounded subset of $\mathbf{R}^n$ meets Γ in finitely many points only. In fact, if $p = \sum_{j=1}^n m_j v_j$ is a lattice point viewed as a column vector, then

$$p = M \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix}$$

where M is the (invertible) matrix having the $v_1, \ldots, v_n$ as columns. We obtain

$$|m_i| \leq \left(\sum_{j=1}^n m_j^2 \right)^{1/2} \leq ||M^{-1}||\, ||p||$$

for all $i = 1, \ldots, n$, where $|| \cdot ||$ denotes the Euclidean norm. Therefore if we require p to lie in a given bounded subset of $\mathbf{R}^n$, then there are only finitely many possibilities for the integers m_j, and thus only finitely many such lattice points. Note that discreteness of Γ implies that Γ, and thus any equivalence class $p + \Gamma$, is a closed subset of $\mathbf{R}^n$.

Equip M with the quotient topology induced by the canonical projection $\pi : \mathbf{R}^n \to M$ that maps each $p \in \mathbf{R}^n$ to its equivalence class $[p] = p + \Gamma$. Then π is continuous. It follows that M is compact since it coincides with the image of $\{ \sum_{j=1}^n x_j v_j \mid 0 \leq x_j \leq 1 \}$ under the projection π. Moreover, π is an open map, as for an open subset W of $\mathbf{R}^n$ we have that $\pi^{-1}(\pi(W)) = \cup_{\gamma \in \Gamma} (W + \gamma)$ is a union of open sets and thus open. It follows that the projection of a countable basis of open sets of $\mathbf{R}^n$ is a countable basis of open sets of M. We also see that the quotient topology is Hausdorff. In fact, given $[p]$, $[q] \in \mathbf{R}^n/\Gamma$, $[p] \neq [q]$, the minimal distance r_{pq} from p to a point in the closed subset $q + \Gamma$ is positive. Let W_p, W_q be the balls of radius $\frac{r_{pq}}{2}$ centered at p, q, respectively. A point $x \in W_p \cap (W_q + \Gamma)$ satisfies $d(x,p) < \frac{r}{2}$ and $d(x, q+\gamma) < \frac{r}{2}$ for some $\gamma \in \Gamma$, and therefore $d(p, q+\gamma) \leq d(p,x) + d(x, q+\gamma) < r$ leading to a contradiction. It follows that $W_p \cap (W_q + \Gamma) = \varnothing$ and hence $\pi(W_p)$, $\pi(W_q)$ are disjoint open neighborhoods of $[p]$, $[q]$, respectively.

We next check that $\pi : \mathbf{R}^n \to M$ is a covering. In fact, discreteness of Γ implies that the minimal distance s from a non-zero lattice point to the origin is positive. Note that s is also the minimal distance from any given point $p \in \mathbf{R}^n$ to another point in $p + \Gamma$. Let V be the ball of radius $\frac{s}{2}$ centered at p. Then $V \cap (V + \gamma) = \varnothing$ 0for all $\gamma \in \Gamma \setminus \{0\}$. Note also that $\pi : V \to \pi(V)$ is continuous, open and injective, thus a homemorphism. Now $\pi^{-1}(\pi(V)) = \cup_{\gamma \in \Gamma}(V + \gamma)$ is a disjoint union of open sets on each of which π is a homeomorphism onto $\pi(V)$, proving that $\pi(V)$ is an evenly covered neighborhood and hence π is a covering map. Since $\mathbf{R}^n$ is simply-connected, this is the universal covering and the fundamental group of M is isomorphic to Γ.

Now we have natural local charts for M defined on any evenly covered neighborhood $U = \pi(V)$ as above. Indeed, write $\pi^{-1}U = \cup_{\gamma\in\Gamma}(V+\gamma)$ and take as chart $\varphi_V = (\pi|_V)^{-1} : U \to V$. If $U' = \pi(V')$ is another evenly covered neighborhood as above with $U \cap U' \neq \varnothing$, consider a connected component W of $U \cap U'$, take $p \in V$ such that $[p] \in W$ and note that there is a unique $\gamma \in \Gamma$ such that $p + \gamma \in V'$. Now $\tau_\gamma \circ \varphi_V|_W$ and $\varphi_{V'}|_W$, where τ_γ denotes the translation by γ, are both lifts of the identity map of $\pi(W)$ and coincide on $[p]$, hence $\tau_\gamma \circ \varphi_V|_W = \varphi_{V'}|_W$ (Theorem 0.2.12). This proves that the transition map $\varphi_{V'} \circ \varphi_V^{-1}$ coincides with τ_γ on W and is thus smooth. In this way we have defined a smooth atlas for M. The covering map $\pi : \mathbf{R}^n \to M$ is smooth and in fact a local diffeomorphism because $\pi|_V$ composed with φ_V on the left yields as local representation the identity, so we indeed have a smooth covering. The smooth structure on M is the unique one that makes $\pi : \mathbf{R}^n \to M$ into a smooth covering (this is more than a covering whose covering map is smooth, compare page 8!).

The transition maps of the above atlas are restrictions of translations of $\mathbf{R}^n$ and thus isometries. In account of this, M acquires a natural quotient Riemannian metric g_Γ, which is the unique one making the covering map π into a local isometry. In fact this requirement implies uniqueness of g_Γ, as it imposes that on an evenly covered neighborhood $U = \pi(V)$ as above, the local chart $\varphi_V = (\pi|_V)^{-1}$ must be a local isometry and so $g_\Gamma = \varphi_V^* g$ on U, where g denotes the canonical metric in $\mathbf{R}^n$. To have existence of g_Γ, we need to check that it is well defined, namely, for another evenly covered neighborhood $U' = \pi(V')$ as above with $U \cap U' \neq \varnothing$ it holds that $\varphi_V^* g = \varphi_{V'}^* g$ on $U \cap U'$. However, this follows from $\varphi_{V'}^* g = \left((\varphi_{V'}\varphi_V^{-1})\varphi_V\right)^* g = \varphi_V^*(\varphi_{V'}\varphi_V^{-1})^* g = \varphi_V^* g$ as $(\varphi_{V'}\varphi_V^{-1})^* g = g$. Note that g_Γ is a flat metric.

As a smooth manifold, M is diffeomorphic to the n-torus T^n. In fact, define a map $f : \mathbf{R}^n \to T^n$ by setting

$$f\Big(\sum_{j=1}^n x_j v_j\Big) = (e^{2\pi i x_1}, \ldots, e^{2\pi i x_n}),$$

where we view S^1 as the set of unit complex numbers. Then f is constant on Γ, so it induces a bijection $\bar f : M \to T^n$. Suitable restrictions of

$$(e^{2\pi i x_1}, \ldots, e^{2\pi i x_n}) \mapsto (x_1, \ldots, x_n)$$

define local charts of T^n whose domains cover it. Now $f = \bar f \circ \pi$ composed on the left with such charts of T^n give $\sum_{j=1}^n x_j v_j \mapsto (x_1, \ldots, x_n)$, the restriction of an invertible linear map. It follows that $\bar f$ is a local diffeomorphism and hence a diffeomorphism.

We remark that different lattices may give rise to nonisometric flat tori, although they will always be locally isometric one to the other since they are all isometrically covered by Euclidean space; in other words, for two given lattices Γ, Γ', suitable restrictions of the identity map id : $\mathbf{R}^n \to \mathbf{R}^n$ induce locally defined isometries $\mathbf{R}^n/\Gamma \to \mathbf{R}^n/\Gamma'$.

One way to globally distinguish the isometry classes of tori obtained from different lattices is to show that they have different isometry groups. To fix ideas, let $n = 2$, and consider in $\mathbf{R}^2$ the lattices Γ, Γ' respectively generated by the bases $\{(1,0),(0,1)\}$ and $\{(1,0),(\frac{1}{2},\frac{\sqrt{3}}{2})\}$. Then $\mathbf{R}^2/\Gamma$ is called a square flat torus and $\mathbf{R}^2/\Gamma'$ is called an hexagonal flat torus. The isotropy subgroup of the square torus at an arbitrary point is isomorphic to the dihedral group D_4 (of order 8) whereas the isotropy subgroup of the hexagonal torus at an arbitrary point is isomorphic to the dihedral group D_3. Hence $\mathbf{R}^2/\Gamma$ and $\mathbf{R}^2/\Gamma'$ are not isometric. See exercise 9 for a characterization of isometric flat tori.

We finish the discussion of this example by noting that we could have introduced the smooth structure on M and the smooth covering $\pi : \mathbf{R}^n \to M$ by invoking Theorem 0.2.13, which we have avoided only for pedagogical reasons. In fact, the elements of Γ can be identified with the translations of $\mathbf{R}^n$ that they define and, in this way, Γ becomes a discrete group acting on $\mathbf{R}^n$. Plainly, the action is free. It is also proper, as this follows from the existence of $r > 0$ such that $d(p, q+\Gamma) \geq r$ if $p \neq q$ and $d(p, p+\Gamma \setminus \{0\}) \geq r$, which was shown above. In the next subsection, we follow and extend this alternative approach to incorporate the construction of the quotient metric.

Riemannian coverings

A *Riemannian covering* between two Riemannian manifolds is a smooth covering that is also a local isometry. For instance, for a lattice Γ in $\mathbf{R}^n$ the projection $\pi : \mathbf{R}^n \to \mathbf{R}^n/\Gamma$ is a Riemannian covering.

If $\tilde{M}$ is a smooth manifold and Γ is a discrete group acting freely and properly by diffeomorphisms on $\tilde{M}$, then the quotient space $M = \Gamma\backslash\tilde{M}$ endowed with the quotient topology admits a unique structure of smooth manifold such that the projection $\pi : \tilde{M} \to M$ is a smooth covering, owing to Theorem 0.2.13. If we assume, in addition, that $\tilde{M}$ is equipped with a Riemannian metric $\tilde{g}$ and Γ acts on $\tilde{M}$ by isometries, then we can show that there is a unique Riemannian metric g on M, called the *quotient metric*, so that $\pi : (\tilde{M}, \tilde{g}) \to (M, g)$ becomes a Riemannian covering, as follows. Around any point $p \in M$, there is an evenly covered neighborhood U such that $\pi^{-1}U = \cup_{i \in I}\tilde{U}_i$. If π is to be a local isometry, we must have

$$g = \left((\pi|_{\tilde{U}_i})^{-1}\right)^* \tilde{g}$$

on U, for any $i \in I$. In more pedestrian terms, we are forced to have

$$g_q(u, v) = \tilde{g}_{\tilde{q}_i}((d\pi_{\tilde{q}_i})^{-1}(u), (d\pi_{\tilde{q}_i})^{-1}(v)), \tag{1.3.2}$$

for all $q \in U$, u, $v \in T_qM$, $i \in I$, where $\tilde{q}_i = (\pi|_{\tilde{U}_i})^{-1}(q)$ is the unique point in the fiber $\pi^{-1}(q)$ that lies in $\tilde{U}_i$. We claim that this definition of g_q does not depend on the choice of point in $\pi^{-1}(q)$. In fact, if $\tilde{q}_j$ is another point in $\pi^{-1}(q)$, there is a unique $\gamma \in \Gamma$ such that $\gamma(\tilde{q}_i) = \tilde{q}_j$. Since $\pi \circ \gamma = \pi$, the chain rule gives that $d\pi_{\tilde{q}_j} \circ d\gamma_{\tilde{q}_i} = d\pi_{\tilde{q}_i}$, so

$$\begin{aligned}\tilde{g}_{\tilde{q}_i}((d\pi_{\tilde{q}_i})^{-1}(u), (d\pi_{\tilde{q}_i})^{-1}(v)) &= \tilde{g}_{\tilde{q}_i}((d\gamma_{\tilde{q}_i})^{-1}(d\pi_{\tilde{q}_j})^{-1}(u), (d\gamma_{\tilde{q}_i})^{-1}(d\pi_{\tilde{q}_j})^{-1}(v)) \\ &= \tilde{g}_{\tilde{q}_j}((d\pi_{\tilde{q}_j})^{-1}(u), (d\pi_{\tilde{q}_j})^{-1}(v)),\end{aligned}$$

since $d\gamma_{\tilde{q}_i} : T_{\tilde{q}_i}\tilde{M} \to T_{\tilde{q}_j}\tilde{M}$ is a linear isometry, checking the claim. Note that g is smooth since it is locally given as a pull-back metric.

On the other hand, if we start with a Riemannian manifold (M, g) and a smooth covering $\pi : \tilde{M} \to M$, then π is in particular an immersion, so we can endow $\tilde{M}$ with the pulled-back metric $\tilde{g}$ and $\pi : (\tilde{M}, \tilde{g}) \to (M, g)$ becomes a Riemannian covering. Let Γ denote the group of deck transformations of $\pi : \tilde{M} \to M$. An element $\gamma \in \Gamma$ satisfies $\pi \circ \gamma = \pi$. Since π is a local isometry, we have that γ is a local isometry, and being a bijection, it must be a global isometry. Hence the group Γ consists of isometries of $\tilde{M}$. If we assume, in addition, that $\pi : \tilde{M} \to M$ is a regular covering (meaning that Γ acts transitively on each fiber of π; this is true, for instance, if $\pi : \tilde{M} \to M$ is the universal covering), then M is diffeomorphic to the orbit space $\Gamma\backslash\tilde{M}$, and since we already know that $\pi : (\tilde{M}, \tilde{g}) \to (M, g)$ is a Riemannian covering, it follows from the uniqueness result of the previous paragraph that g must be the quotient metric of $\tilde{g}$.

The real projective space $\mathbf{R}P^n$

As a set, $\mathbf{R}P^n$ is the set of all lines through the origin in $\mathbf{R}^{n+1}$. It can also be naturally viewed as a quotient space in two ways. In the first one, we define an equivalence relation among points in $\mathbf{R}^{n+1} \setminus \{0\}$ by declaring x and y to be equivalent if they lie in the same line, namely, if there exists $\lambda \in \mathbf{R} \setminus \{0\}$ such that $y = \lambda x$. In the second one, we simply note that every line meets the unit sphere in $\mathbf{R}^{n+1}$ in two antipodal points, so we can also view $\mathbf{R}P^n$ as a quotient space of S^n and, in this case, $x, y \in S^n$ are equivalent if and only if $y = \pm x$. Of course, in both cases $\mathbf{R}P^n$ acquires the same quotient topology.

Next, we reformulate our point of view slightly by introducing the group Γ consisting of two isometries of S^n, namely the identity map and the antipodal map. Then Γ obviously acts freely and properly (it is a finite group!) on S^n, and the resulting quotient smooth structure makes $\mathbf{R}P^n$ into a smooth manifold. Furthermore, as the action of Γ is also isometric, $\mathbf{R}P^n$ immediately acquires a Riemannian metric such that $\pi : S^n \to \mathbf{R}P^n$ is a Riemannian covering.

The Klein bottle

Let $\tilde{M} = \mathbf{R}^2$, let $\{v_1, v_2\}$ be a basis of $\mathbf{R}^2$, and let Γ be the discrete group of transformations of $\mathbf{R}^2$ generated by the affine linear maps

$$\gamma_1(x_1 v_1 + x_2 v_2) = \left(x_1 + \frac{1}{2}\right) v_1 - x_2 v_2 \quad \text{and} \quad \gamma_2(x_1 v_1 + x_2 v_2) = x_1 v_1 + (x_2 + 1) v_2.$$

It is easy to see that Γ acts freely and properly on $\mathbf{R}^2$, so we get a quotient manifold $\mathbf{R}^2/\Gamma$ which is called the *Klein bottle* K^2. It is a compact non-orientable manifold, since γ_1 reverses the orientation of $\mathbf{R}^2$. It follows that K^2 cannot be embedded in $\mathbf{R}^3$ by the Jordan-Brouwer separation theorem; however, it is easy to see that it can be immersed there.

Consider $\mathbf{R}^2$ equipped with its canonical metric. Note that γ_2 is always an isometry of $\mathbf{R}^2$, but so is γ_1 if and only if the basis $\{v_1, v_2\}$ is orthogonal. In this case, Γ acts by isometries on $\mathbf{R}^2$ and K^2 inherits a flat metric so that the projection $\mathbf{R}^2 \to K^2$ is a Riemannian covering.

Riemannian submersions

Let $\pi : M \to N$ be a smooth submersion between two smooth manifolds. Then $\mathcal{V}_p = \ker d\pi_p$ for $p \in M$ defines a smooth distribution on M which is called the *vertical distribution*. Clearly, $\mathcal{V}$ can also be given by the tangent spaces of the fibers of π. In general, there is no canonical choice of a complementary distribution of $\mathcal{V}$ in TM, but in the case in which M comes equipped with a Riemannian metric, one can naturally construct such a complement $\mathcal{H}$ by setting $\mathcal{H}_p$ to be the orthogonal complement of $\mathcal{V}_p$ in T_pM. Then $\mathcal{H}$ is a smooth distribution which is called the *horizontal distribution*. Note that $d\pi_p$ induces an isomorphism between $\mathcal{H}_p$ and $T_{\pi(p)}N$ for every $p \in M$.

Having these preliminary remarks at hand, we can now define a smooth submersion $\pi : (M, g) \to (N, h)$ between two Riemannian manifolds to be a *Riemannian submersion* if $d\pi_p$ induces an isometry between $\mathcal{H}_p$ and $T_{\pi(p)}N$ for every $p \in M$. Note that Riemannian coverings are particular cases of Riemannian submersions.

Let (M, g) and (N, h) be Riemannian manifolds. A quite trivial example of a Riemannian submersion is the projection $(M \times N, g + h) \to (M, g)$ (or $(M \times N, g + h) \to (N, h)$). More generally, if f is a nowhere zero smooth function on N, the projection from $(M \times N, f^2 g + h)$ onto (N, h) is a Riemannian submersion. In this case, the fibers of the submersion are homothetic but not necessarily isometric one to the other. A Riemannian manifold of the form $(M \times N, f^2 g + h)$ is called a *warped product*.

Recall that if $\tilde{M}$ is a smooth manifold and G is a Lie group acting freely and properly on $\tilde{M}$, then the quotient space $M = G\backslash\tilde{M}$ endowed with the quotient topology admits a unique structure of smooth manifold such that the projection $\pi : \tilde{M} \to M$ is a (surjective) submersion (Theorem 0.4.16). If in addition we assume that $\tilde{M}$ is equipped with a Riemannian metric $\tilde{g}$ and G acts on $\tilde{M}$ by isometries, then we can show that there is a unique Riemannian metric g on M, called the *quotient metric*, so that $\pi : (\tilde{M}, \tilde{g}) \to (M, g)$ becomes a Riemannian submersion. Indeed, given a point $p \in M$ and tangent vectors u, $v \in T_pM$, we set

$$g_p(u, v) = \tilde{g}_{\tilde{p}}(\tilde{u}, \tilde{v}), \tag{1.3.3}$$

where $\tilde{p}$ is any point in the fiber $\pi^{-1}(p)$ and $\tilde{u}$, $\tilde{v}$ are the unique vectors in $\mathcal{H}_{\tilde{p}}$ satisfying $d\pi_{\tilde{p}}(\tilde{u}) = u$ and $d\pi_{\tilde{p}}(\tilde{v}) = v$. The proof that $\tilde{g}$ is well defined is similar to the proof that the quotient metric is well defined in the case of a Riemannian covering, namely, choosing a different point $\tilde{p}' \in \pi^{-1}(p)$, one has unique vectors $\tilde{u}'$, $\tilde{v}' \in \mathcal{H}_{\tilde{p}'}$ that project to u, v, but $\tilde{g}_{\tilde{p}'}(\tilde{u}', \tilde{v}')$ gives the same result as above because $\tilde{p}' = \Phi(g)\tilde{p}$ for some $g \in G$, $d(\Phi(g))_{\tilde{p}} : \mathcal{H}_{\tilde{p}} \to \mathcal{H}_{\tilde{p}'}$ is an isometry and maps $\tilde{u}$, $\tilde{v}$ to $\tilde{u}'$, $\tilde{v}'$ respectively. The proof that $\tilde{g}$ is smooth is also similar, but needs an extra ingredient. Let $P_{\tilde{p}} : T_{\tilde{p}}\tilde{M} \to \mathcal{H}_{\tilde{p}}$ denote the orthogonal projection. It is known that $\pi : \tilde{M} \to M$ admits local sections, so let $s : U \to \tilde{M}$ be a local section defined on an open set U of M. Now we can rewrite (1.3.3) as

$$g_q(u, v) = \tilde{g}_{s(q)}(P_{s(q)} ds_q(u), P_{s(q)} ds_q(v)),$$

where $q \in U$. Since $\mathcal{V}$ as a distribution is locally defined by smooth vector fields, it is easy to check that P takes locally defined smooth vector fields on TM to locally defined smooth vector fields on TM. It follows that g is smooth. Finally, the requirement that π be a Riemannian submersion forces g to be given by formula (1.3.3), and this shows the uniqueness of g.

The complex projective space $\mathbf{C}P^n$

The definition of $\mathbf{C}P^n$ is similar to that of $\mathbf{R}P^n$ in that we replace real numbers by complex numbers. Namely, as a set, $\mathbf{C}P^n$ is the set of all complex lines through the origin in $\mathbf{C}^{n+1}$, so it can be viewed as the quotient of $\mathbf{C}^{n+1} \setminus \{0\}$ by the multiplicative group $\mathbf{C} \setminus \{0\}$ as well as the quotient of the unit sphere S^{2n+1} of $\mathbf{C}^{n+1}$ (via its canonical identification with $\mathbf{R}^{2n+2}$) by the multiplicative group of unit complex numbers S^1. Here the action of S^1 on S^{2n+1} is given by multiplication of the coordinates (since $\mathbf{C}$ is commutative, it is unimportant whether S^1 multiplies on the left or on the right). This action is clearly free and it is also proper since S^1 is compact. Further, the multiplication $L_z : S^{2n+1} \to S^{2n+1}$ by a unit complex number $z \in S^1$ is an isometry. In fact, S^{2n+1} has the induced metric from $\mathbf{R}^{2n+2}$, the Euclidean scalar product is the real part of the Hermitian inner product $(\cdot, \cdot)$ of $\mathbf{C}^{n+1}$ and $(L_z x, L_z y) = (zx, zy) = ||z||^2(x, y) = (x, y)$ for all x, $y \in \mathbf{C}^{n+1}$. It follows that $\mathbf{C}P^n = S^{2n+1}/S^1$ has the structure of a compact smooth manifold of dimension $2n$. Moreover there is a natural Riemannian metric which makes the projection $\pi : S^{2n+1} \to \mathbf{C}P^n$ into a Riemannian submersion. This quotient metric is classically called the *Fubini-Study metric* on $\mathbf{C}P^n$.

We want to explicitly construct the smooth structure on $\mathbf{C}P^n$ and prove that $\pi : S^{2n+1} \to \mathbf{C}P^n$ is a submersion in order to better familiarize ourselves with such an important example. For each $p \in \mathbf{C}P^n$, we construct a local chart around p. View p as a one-dimensional subspace of $\mathbf{C}^{n+1}$ and denote its Hermitian orthogonal complement by $p^\perp$. The subset of all lines which are not parallel to $p^\perp$ is an open subset of $\mathbf{C}P^n$, which we denote by $\mathbf{C}P^n \setminus p^\perp$. Fix a unit vector $\tilde p$ lying in the line p. The local chart is

$$\varphi^p : \mathbf{C}P^n \setminus p^\perp \to p^\perp, \qquad q \mapsto \frac{1}{(\tilde q, \tilde p)}\tilde q - \tilde p,$$

where $\tilde q$ is any nonzero vector lying in q. In other words, q meets the affine hyperplane $\tilde p + p^\perp$ at a unique point $\frac{1}{(\tilde q,\tilde p)}\tilde q$ which we orthogonally project to $p^\perp$ to get $\varphi^p(q)$. (Note that $p^\perp$ can be identified with $\mathbf{R}^{2n}$ simply by choosing a basis.) The inverse of φ^p is the map that takes $v \in p^\perp$ to the line through $\tilde p + v$. Therefore, for $p' \in \mathbf{C}P^n$, we see that the transition map $\varphi^{p'} \circ (\varphi^p)^{-1} : \{\, v \in p^\perp \mid v + \tilde p \notin p'^\perp \,\} \to \{\, v' \in p'^\perp \mid v' + \tilde p' \notin p^\perp \,\}$ is given by

$$v \mapsto \frac{1}{(v + \tilde p, \tilde p')}(v + \tilde p) - \tilde p', \tag{1.3.4}$$

and hence smooth.

Next we prove that the projection $\pi : S^{2n+1} \to \mathbf{C}P^n$ is a smooth submersion. Let $\tilde p \in S^{2n+1}$. Since the fibers of π are just the S^1-orbits, the vertical space $\mathcal{V}_{\tilde p} = \mathbf{R}(i\tilde p)$. It follows that the horizontal space $\mathcal{H}_{\tilde p} \subset T_{\tilde p}S^{2n+1}$ is the Euclidean orthogonal complement of $\mathbf{R}\{\tilde p, i\tilde p\} = \mathbf{C}\tilde p$ in $\mathbf{C}^{2n+1}$, namely, $p^\perp$ where $p = \pi(\tilde p)$. It suffices to check that $d\pi_{\tilde p}$ is an isomorphism from $\mathcal{H}_{\tilde p}$ onto $T_p\mathbf{C}P^n$, or, $d(\varphi^p \circ \pi)_{\tilde p}$ is an isomorphism from $p^\perp$ to itself. Let v be a unit vector in $p^\perp$. Then $t \mapsto \cos t\, \tilde p + \sin t\, v$ is a curve in S^{2n+1} with initial point $\tilde p$ and initial speed v, so using that $(\cos t\, \tilde p + \sin t\, v, \tilde p) = \cos t$ we have

$$\begin{aligned} d(\varphi^p \circ \pi)_{\tilde p}(v) &= \frac{d}{dt}\Big|_{t=0} (\varphi^p \circ \pi)(\cos t\, \tilde p + \sin t\, v) \\ &= \frac{d}{dt}\Big|_{t=0} \frac{1}{\cos t}(\cos t\, \tilde p + \sin t\, v) - \tilde p \\ &= v, \end{aligned}$$

completing the check.

One-dimensional Riemannian manifolds

Let (M, g) be a Riemannian manifold and let $\gamma : [a, b] \to M$ be a piecewise C^1 curve. Then the *length* of γ is defined to be

$$L(\gamma) = \int_a^b g_{\gamma(t)}(\gamma'(t), \gamma'(t))^{1/2}\, dt. \tag{1.3.5}$$

It is easily seen that the length of a curve does not change under re-parametrization. Moreover, every regular curve (i.e. satisfying $\gamma'(t) \neq 0$ for all t) admits a natural parametrization given by arc-length. Namely, let

$$s(t) = \int_a^t g_{\gamma(\tau)}(\gamma'(\tau), \gamma'(\tau))^{1/2}\, d\tau.$$

Then $\dfrac{ds}{dt} = g_{\gamma(t)}(\gamma'(t), \gamma'(t))^{1/2}(t) > 0$, so s can be taken as a new parameter, and then

$$L(\gamma|_{[0,s]}) = s$$

and

$$(\gamma^* g)_t = g_{\gamma(t)}(\gamma'(t), \gamma'(t))dt^2 = ds^2. \tag{1.3.6}$$

Suppose now that (M, g) is a one-dimensional Riemannian manifold. Then any connected component of M is diffeomorphic either to $\mathbf{R}$ or to S^1. In any case, a neighborhood of any point $p \in M$ can be viewed as a regular smooth curve in M and, in a parametrization by arc-length, the local expression of the metric g is the same, namely, given by (1.3.6). It follows that all the one-dimensional Riemannian manifolds are locally isometric among themselves.

Lie groups ★

The natural class of Riemannian metrics to be considered in Lie groups is the class of Riemannian metrics that possess some kind of invariance, be it left, right or both. Let G be a Lie group. A *left-invariant Riemannian metric* on G is a Riemannian metric with respect to which the left translations of G are isometries. Similarly, a *right-invariant Riemannian metric* is defined. A Riemannian metric on G that is both left- and right-invariant is called a *bi-invariant Riemannian metric.*

Left-invariant Riemannian metrics (henceforth, left-invariant metrics) are easy to construct on any given Lie group G. In fact, given any inner product $\langle , \rangle$ in its Lie algebra $\mathfrak{g}$, which we identify with the tangent space at the identity T_1G, one sets $g_1 = \langle , \rangle$ and uses the left translations to pull back g_1 to the other tangent spaces, namely one sets

$$g_x(u, v) = g_1\big(d(L_{x^{-1}})_x(u) \, , \, d(L_{x^{-1}})_x(v) \big),$$

where $x \in G$ and u, $v \in T_xG$. This defines a smooth Riemannian metric, since $g(X, Y)$ is constant (and hence smooth) for any pair (X, Y) of left-invariant vector fields, and any smooth vector field on G is a linear combination of left-invariant vector fields with smooth functions as coefficients. By the very construction of g, the $d(L_x)_1$ for $x \in G$ are linear isometries, so the composition of linear isometries $d(L_x)_y = d(L_{xy})_1 \circ d(L_y)_1^{-1}$ is also a linear isometry for x, $y \in G$. This checks that all the left-translations are isometries and hence that g is left-invariant. (Equivalently, one can define g by choosing a global frame of left-invariant vector fields on G and declaring it to be orthonormal at every point of G.) It follows that the set of left-invariant metrics in G is in bijection with the set of inner products on $\mathfrak{g}$. Of course, similar remarks apply to right-invariant metrics.

Bi-invariant metrics are more difficult to come up with. Starting with a fixed left-invariant metric g on G, we want to find conditions for g to be also right-invariant. Reasoning similarly as in the previous paragraph, we see that it is necessary and sufficient that the $d(R_x)_1$ for $x \in G$ be linear isometries. Further, by differentiating the obvious identity $R_x = L_x \circ \mathrm{Inn}(x^{-1})$ at 1, we get that

$$d(R_x)_1 = d(L_x)_1 \circ \mathrm{Ad}(x^{-1})$$

for $x \in G$. From this identity, we get that g is right-invariant if and only if the $\mathrm{Ad}(x) : \mathfrak{g} \to \mathfrak{g}$ for $x \in G$ are linear isometries with respect to $\langle , \rangle = g_1$. In this case, $\langle , \rangle$ is called an Ad-*invariant inner product* on $\mathfrak{g}$.

In view of the previous discussion, applying the following proposition to the adjoint representation of a compact Lie group on its Lie algebra yields that *any compact Lie group admits a bi-invariant Riemannian metric.*

1.3.7 Proposition *Let $\rho : G \to \mathbf{GL}(V)$ be a representation of a Lie group on a real vector space V such that the closure $\rho(G)$ is relatively compact in $\mathbf{GL}(V)$. Then there exists an inner product $\langle , \rangle$ on V with respect to which the $\rho(x)$ for $x \in G$ are orthogonal transformations.*

Proof. Let $\tilde{G}$ denote the closure of $\rho(G)$ in $\mathbf{GL}(V)$. Then ρ factors through the inclusion $\tilde{\rho} : \tilde{G} \to \mathbf{GL}(V)$ and it suffices to prove the result for $\tilde{\rho}$ instead of ρ. By assumption, $\tilde{G}$ is compact, so without loss of generality we may assume in the following that G is compact.

Let $\langle , \rangle_0$ be any inner product on V and fix a right-invariant Haar measure dx on G. Set

$$\langle u, v \rangle = \int_G \langle \rho(x)u, \rho(x)v \rangle_0 \, dx,$$

where u, $v \in V$. It is easy to see that this defines a positive-definite bilinear symmetric form $\langle , \rangle$ on V. Moreover, if $y \in G$, then

$$\begin{aligned} \langle \rho(y)u, \rho(y)v \rangle &= \int_G \langle \rho(x)\rho(y)u, \rho(x)\rho(y)v \rangle_0 \, dx \\ &= \int_G \langle \rho(xy)u, \rho(xy)v \rangle_0 \, dx \\ &= \langle u, v \rangle, \end{aligned}$$

where in the last equality we have used that dx is right-invariant. Note that we have used the compactness of G only to guarantee that the above integrands have compact support. □

In later chapters, we will explain the special properties that bi-invariant metrics on Lie groups have.

Homogeneous spaces ★

It is apparent that for a generic Riemannian manifold (M, g), the isometry group $\mathrm{Isom}(M, g)$ is trivial. Indeed, Riemannian manifolds with large isometry groups have a good deal of symmetries. In particular, in the case in which $\mathrm{Isom}(M, g)$ is transitive on M, (M, g) is called a *Riemannian homogeneous space* or a *homogeneous Riemannian manifold*. Explicitly, this means that given any two points of M there exists an isometry of M that maps one point to the other. In this case, of course it may happen that a subgroup of $\mathrm{Isom}(M, g)$ is already transitive on M.

Let (M, g) be a homogeneous Riemannian manifold, and let G be a subgroup of $\mathrm{Isom}(M, g)$ acting transitively on M. Then the isotropy subgroup H at an arbitrary fixed point $p \in M$ is compact and M is diffeomorphic to the quotient space G/H. In this case, we also say that the Riemannian metric g on M is *G-invariant*.

Recall that if G is a Lie group and H is a closed subgroup of G, then there exists a unique structure of smooth manifold on the quotient G/H such that the projection $G \to G/H$ is a submersion and the action of G on G/H by left translations is smooth. (Theorem 0.4.18). A manifold of the form G/H is called a homogeneous space. In some cases, one can also start with a homogeneous space G/H and construct G-invariant metrics on G/H. For instance, if G is equipped with a left-invariant metric *that is also right-invariant with respect to H*, then it follows that the quotient G/H inherits a quotient Riemannian metric such that the projection $G \to G/H$ is a Riemannian submersion and the action of G on G/H by left translations is isometric. In this way, G/H becomes a Riemannian homogeneous space. A particular, important case of this construction is when the Riemannian metric on G that we start with is bi-invariant; in this case, G/H is called a *normal homogeneous space*. In general, a homogeneous space G/H for arbitrary G, H may admit several distinct G-invariant Riemannian metrics, or may admit no such metrics at all.

Let $M = G/H$ be a homogeneous space, where H is the isotropy subgroup at $p \in M$. Then the *isotropy representation* at p is the homomorphism

$$H \to O(T_pM), \qquad h \mapsto dh_p. \tag{1.3.8}$$

1.3.9 Lemma *The isotropy representation of G/H at p is equivalent to the adjoint representation of H on $\mathfrak{g}/\mathfrak{h}$.*

1.3.10 Proposition *a. There exists a G-invariant Riemannian metric on G/H if and only if the image of the adjoint representation of H on $\mathfrak{g}/\mathfrak{h}$ is relatively compact in $GL(\mathfrak{g}/\mathfrak{k})$.*

b. In case the condition in (a) is true, the G-invariant metrics on G/H are in bijective correspondence with the $\mathrm{Ad}_G(H)$-invariant inner products on $\mathfrak{g}/\mathfrak{h}$.

1.4 Exercises

1 Show that the Riemannian product of $(0, +\infty)$ and S^{n-1} is isometric to the cylinder

$$C = \{ (x_0, \ldots, x_n) \in \mathbf{R}^{n+1} \mid x_1^2 + \cdots + x_n^2 = 1 \quad \text{and} \quad x_0 > 0 \}.$$

2 The *catenoid* is the surface of revolution in $\mathbf{R}^3$ with the z-axis as axis of revolution and the catenary $x = \cosh z$ in the xz-plane as generating curve. The *helicoid* is the ruled surface in $\mathbf{R}^3$ consisting of all the lines parallel to the xy plane that intersect the z-axis and the helix $t \mapsto (\cos t, \sin t, t)$.

a. Write natural parametrizations for the catenoid and the helicoid.

b. Consider the catenoid and the helicoid with the metrics induced from $\mathbf{R}^3$, and find the local expressions of these metrics with respect to the parametrizations in item (a).

c. Show that the local expressions in item (b) coincide, possibly up to a change of coordinates, and deduce that the catenoid and the helicoid are locally isometric.

d. Show that the catenoid and the helicoid cannot be isometric because of their topology.

3 Consider the real hyperbolic space $(\mathbf{R}H^n, g)$ as defined in section 1.3. Let $\mathbf{B}^n$ be the open unit ball of $\mathbf{R}^n$ embedded in $\mathbf{R}^{n+1}$ as

$$\mathbf{B}^n = \{ (x_0, \ldots, x_n) \in \mathbf{R}^{n+1} \mid x_0 = 0 \quad \text{and} \quad x_1^2 + \cdots + x_n^2 < 1 \}.$$

Define a map $f : \mathbf{R}H^n \to \mathbf{B}^n$ by setting $f(x)$ to be the unique point of $\mathbf{B}^n$ lying in the line joining $x \in \mathbf{R}H^n$ and the point $(-1, 0, \ldots, 0) \in \mathbf{R}^{n+1}$. Prove that f is a diffeomorphism and, setting $g_1 = (f^{-1})^* g$, we have that

$$g_1|_x = \frac{4}{(1 - \langle x, x \rangle)^2} \left(dx_1^2 + \cdots + dx_n^2 \right),$$

where $x = (0, x_1, \ldots, x_n) \in \mathbf{B}^n$. Deduce that $\mathbf{R}H^n$ is conformally flat.

$(\mathbf{B}^n, g_1)$ is called the *Poincaré ball model* of $\mathbf{R}H^n$.

4 Consider the open unit ball $\mathbf{B}^n = \{ (x_1, \ldots, x_n) \in \mathbf{R}^n \mid x_1^2 + \cdots x_n^2 < 1 \}$ equipped with the metric g_1 as in Exercise 3. Prove that the inversion of $\mathbf{R}^n$ on the sphere of center $(-1, 0, \ldots, 0)$ and radius $\sqrt{2}$ defines a diffeomorphism f_1 from $\mathbf{B}^n$ onto the upper half-space

$$\mathbf{R}^n_+ = \{ (x_1, \ldots, x_n) \in \mathbf{R}^n \mid x_1 > 0 \},$$

and that the metric $g_2 = (f_1^{-1})^* g_1$ is given by

$$g_2|_x = \frac{1}{x_1^2} \left(dx_1^2 + \cdots + dx_n^2 \right),$$

where $x = (x_1, \ldots, x_n) \in \mathbf{R}^n_+$.

$(\mathbf{R}^n_+, g_2)$ is called the *Poincaré upper half-space model* of $\mathbf{R}H^n$.

5 Consider the Poincaré upper half-plane model $\mathbf{R}^2_+ = \{ (x, y) \in \mathbf{R}^2 \mid y > 0 \}$ with the metric $g_2 = \frac{1}{y^2}\left(dx^2 + dy^2\right)$ (case $n = 2$ in Exercise 4).

(i) Check that the following transformations of $\mathbf{R}^2_+$ into itself are isometries:

(a) $\tau_a(x, y) = (x + a, y)$ for $a \in \mathbf{R}$;
(b) $h_r(x, y) = (rx, ry)$ for $r > 0$;
(c) $R(x, y) = \left(\dfrac{x}{x^2 + y^2}, \dfrac{y}{x^2 + y^2}\right)$.

(ii) Deduce from (a) and (b) that $\mathbf{R}^+_2$ is homogeneous.

(iii) In complex notation, the half-plane model of real hyperbolic space is $\mathbf{R}^2_+ = \{ z \in \mathbf{C} \mid \Im z > 0 \}$ with the Riemannian metric $g = \frac{1}{||\Im z||^2} dz d\bar{z}$. Deduce from (a), (b) and (c) that $T(z) = (az + b)/(cz + d)$ for a, b, c, $d \in \mathbf{R}$ with $ad - bc > 0$ defines an isometry of $(\mathbf{R}^2_+, g)$.

(iv) A homogeneous Riemannian manifold $M = G/G_p$ is called *isotropic* if the isotropy group G_p acts transitively on the unit sphere of T_pM via the isotropy representation (1.3.8). Use (iii) to show that $\mathbf{R}^2_+$ is isotropic.

6 Use stereographic projection to prove that S^n is conformally flat.

7 Consider the parametrized curve

$$\begin{cases} x &= t - \tanh t \\ y &= \frac{1}{\cosh t} \end{cases}$$

The surface of revolution in $\mathbf{R}^3$ constructed by revolving it around the x-axis is called the *pseudo-sphere*. Note that the pseudo-sphere is singular along the circle obtained by revolving the point $(0, 1)$.

a. Prove that the pseudo-sphere with the singular circle taken away is locally isometric to the upper half plane model of $\mathbf{R}H^2$.

b. Show that the Gaussian curvature of the pseudo-sphere is -1.

8 Let Γ be the lattice in $\mathbf{R}^n$ defined by the basis $\{v_1, \ldots, v_n\}$, and denote by g_Γ the Riemannian metric that it defines on T^n. Show that in some product chart of $T^n = S^1 \times \cdots \times S^1$ the local expression

$$g_\Gamma = \sum_{i,j} \langle v_i, v_j \rangle \, dx_i \otimes dx_j$$

holds, where $\langle , \rangle$ denotes the standard scalar product in $\mathbf{R}^n$.

9 Let Γ and Γ' be two lattices in $\mathbf{R}^n$, and denote by g_Γ, $g_{\Gamma'}$ the Riemannian metrics that they define on T^n, respectively.

a. Prove that (T^n, g_Γ) is isometric to $(T^n, g_{\Gamma'})$ if and only if there exists an isometry $f : \mathbf{R}^n \to \mathbf{R}^n$ such that $f(\Gamma) = \Gamma'$. (Hint: You may use the result of exercise 2 of chapter 3.)

b. Use part (a) to see that (T^n, g_Γ) is isometric to the Riemannian product of n copies of S^1 if and only if Γ is the lattice associated to an orthonormal basis of $\mathbf{R}^n$.

10 Let Γ be the lattice of $\mathbf{R}^2$ spanned by an orthogonal basis $\{v_1, v_2\}$ and consider the associated rectangular flat torus T^2.

a. Prove that the map γ of $\mathbf{R}^2$ defined by $\gamma(x_1 v_1 + x_2 v_2) = (x_1 + \frac{1}{2})v_1 - x_2 v_2$ induces an isometry of T^2 of order two.

b. Prove that T^2 double covers a Klein bottle K^2.

11 Prove that $\mathbf{R}^n \setminus \{0\}$ is isometric to the warped product $((0,+\infty) \times S^{n-1}, dr^2 + r^2 g)$, where r denotes the coordinate on $(0,+\infty)$ and g denotes the standard Riemannian metric on S^{n-1}.

12 Let G be a Lie group equal to one of $\mathbf{O}(n)$, $\mathbf{U}(n)$ or $\mathbf{SU}(n)$, and denote its Lie algebra by $\mathfrak{g}$. Prove that for any $c > 0$

$$\langle X, Y \rangle = -c \operatorname{trace}(XY),$$

where X, $Y \in \mathfrak{g}$, defines an Ad-invariant inner product on $\mathfrak{g}$.

13 Consider the special unitary group $\mathbf{SU}(2)$ equipped with a bi-invariant metric induced from an Ad-invariant inner product on $\mathfrak{su}(2)$ as in the previous exercise with $c = \frac{1}{2}$. Show that the map

$$\begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \mapsto \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

where α, $\beta \in \mathbf{C}$ and $|\alpha|^2 + |\beta|^2 = 1$, defines an isometry from $\mathbf{SU}(2)$ onto S^3. Here $\mathbf{C}^2$ is identified with $\mathbf{R}^4$ and S^3 is viewed as the unit sphere in $\mathbf{R}^4$.

14 Show that $\mathbf{R}P^1$ equipped with the quotient metric from $S^1(1)$ is isometric to $S^1(\frac{1}{2})$. Show that $\mathbf{C}P^1$ equipped with the Fubini-Study metric is isometric to $S^2(\frac{1}{2})$.

15 (Sylvester's law of inertia) Let $B : V \times V \to \mathbf{R}$ be a symmetric bilinear form on a finite-dimensional real vector space V. For each basis $E = (e_1, \ldots, e_n)$ of V, we associate a symmetric matrix $B_E = (B(e_i, e_j))$.

a. Check that $B(u,v) = v_E^t B_E u_E$ for all u, $v \in V$, where u_E (resp. v_E) denotes the column vector representing the vector u (resp. v) in the basis E.

b. Suppose $F = (f_1, \ldots, f_n)$ is another basis of V such that

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}.$$

for a real matrix A of order n. Show that $B_E = A B_F A^t$.

c. Prove that there exists a basis E of V such that B_E has the form

$$\begin{pmatrix} I_{n-i-k} & 0 & 0 \\ 0 & -I_i & 0 \\ 0 & 0 & 0_k \end{pmatrix},$$

where I_m denotes an identity block of order m, and 0_m denotes a null block of order m.

d. Prove that there is a B-orthogonal decomposition

$$V = V_+ \oplus V_- \oplus V_0$$

where B is positive definite on V_+ and negative definite on V_-, V_0 is the kernel of B (the set of vectors B-orthogonal to V), $i = \dim V_-$ and $k = \dim V_0$. Prove also that i is the maximal dimension of a subspace of V on which B is negative definite. Deduce that i and k are invariants of B. They are respectively called the *index* and *nullity* of B. Of course, B is nondegenerate if and only if $k = 0$, and it is positive definite if and only if $k = i = 0$.

e. Check that the Lorentzian metric of $\mathbf{R}^{1,n}$ restricts to a positive definite symmetric bilinear form on the tangent spaces to the hyperboloid modeling $\mathbf{R}H^n$.

1.5 Additional notes

§1 Riemannian manifolds were defined as abstract smooth manifolds equipped with Riemannian metrics. One class of examples of Riemannian manifolds is of course furnished by the Riemannian submanifolds of Euclidean space. On the other hand, a very deep theorem of Nash [Nas56] states that every abstract Riemannian manifold admits an isometric embedding into Euclidean space, so that it can be viewed as an embedded Riemannian submanifold of Euclidean space. In view of this, one might be tempted to ask why bother to consider abstract Riemannian manifolds in the first place. The reason is that Nash's theorem is an existence result: for a given Riemannian manifold, it does not supply an explicit embedding of it into Euclidean space. Even if an isometric embedding is known, there may be more than one (up to congruence) or there may be no canonical one. Also, an explicit embedding may be too complicated to describe. Finally, a particular embedding is sometimes distracting because it highlights some specific features of the manifold at the expense of some other features, which may be undesirable.

§2 From the point of view of foundations of the theory of smooth manifolds, the following assertions are equivalent for a smooth manifold M whose underlying topological space is assumed to be Hausdorff but not necessarily second-countable:

a. The topology of M is paracompact.

b. M admits smooth partitions of unity.

c. M admits Riemannian metrics.

In fact, as is standard in the theory of smooth manifolds, second-countability of the topology of M (together with the Hausdorff property) implies its paracompactness and this is used to prove the existence of smooth partitions of unity [War83, chapter 1]. Next, Riemannian metrics are constructed on M by using partitions of unity as we did in Proposition 1.2.3. Finally, the underlying topology of a Riemannian manifold is metrizable according to Proposition 3.2.3, and every metric space is paracompact.

§3 The pseudo-sphere constructed in Exercise 7 was introduced by Beltrami [Bel68] in 1868 as a local model for the Lobachevskyan geometry. This means that the geodesic lines and their segments on the pseudo-sphere play the role of straight lines and their segments on the Lobachevsky plane. In 1900, Hilbert posed the question of whether there exists a surface in three-dimensional Euclidean space whose intrinsic geometry coincides completely with the geometry of the Lobachevsky plane. Using a simple reasoning, it follows that if such a surface does exist, it must have constant negative curvature and be complete (see chapter 3 for the notion of completeness).

As early as 1901, Hilbert solved this problem [Hil01] (see also [Hop89, chapter IX]), and in the negative sense, so that no complete surface of constant negative curvature exists in three-dimensional Euclidean space. This theorem has attracted the attention of geometers over a number of decades, and continues to do so today. The reason for this is that a number of interesting questions are related to it and to its proof. For instance, the occurrence of a singular circle on the pseudo-sphere is not coincidental, but is in line with Hilbert's theorem.

Connections

2.1 Introduction

Contemplate $\mathbf{R}^n$. Of course, the presence of the identity map as a global chart allows one to canonically identify the tangent spaces of $\mathbf{R}^n$ at its various points with $\mathbf{R}^n$ itself. Therefore, a smooth vector field X in $\mathbf{R}^n$ can be viewed simply as a smooth map $X : \mathbf{R}^n \to \mathbf{R}^n$. Thus, one has a canonical way of differentiating vector fields in $\mathbf{R}^n$, namely, if X, $Y : \mathbf{R}^n \to \mathbf{R}^n$ are two vector fields, then the derivative of Y along X is the directional derivative $dY(X) = X(Y)$.

Whereas a smooth manifold M comes already equipped with a notion of derivative of smooth maps, there is no canonical way to differentiate vector fields on M. We solve this problem by considering all possible ways of defining derivatives of vector fields. Any such choice is called a connection. The name originates from the fact that, at least along a given curve, a connection provides a way to identify ("connect") tangent spaces of M at different points; this is the idea of parallel transport along the curve. A geodesic is then a curve whose velocity vector is constant in this sense.

The main consequence of the theory of connections for Riemannian geometry is that a Riemannian metric on M uniquely specifies a connection on M, called the Levi-Cività connection. In the case in which M is a surface in $\mathbf{R}^3$, for the Levi-Cività connection on M we recover the derivative in $\mathbf{R}^3$ projected back to M.

Connections can be defined in a variety of ways. We will use the Koszul formalism.

2.2 Connections

Let M be a smooth manifold. A *(Koszul) connection* in M is an $\mathbf{R}$-bilinear map $\nabla : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$, where we write $\nabla_X Y$ instead of $\nabla(X, Y)$, such that

a. $\nabla_{fX} Y = f\nabla_X Y$, and

b. $\nabla_X(fY) = X(f)Y + f\nabla_X Y$ (Leibniz rule)

for every X, $Y \in \Gamma(TM)$ and $f \in C^\infty(M)$.

Let ∇ be a connection in a smooth manifold M. We want to analyse the dependence of ∇ on its arguments. To begin with, we claim that, for a given open set U in M, $(\nabla_X Y)|_U$ depends only on $X|_U$ and $Y|_U$. Indeed, let X', $Y' \in \Gamma(TM)$ be vector fields satisfying $X'|_U = X|_U$ and $Y'|_U = Y|_U$. Fix $p \in U$. Construct a smooth function f on M with support contained in U and such that $f \equiv 1$ on some neighborhood V of p with $V \subset \bar{V} \subset U$. Then, using part (a) in the definition of connection and the fact that $fX = fX'$ on M,

$$(\nabla_X Y)_p = f(p)(\nabla_X Y)_p = (f\nabla_X Y)_p = (\nabla_{fX} Y)_p = (\nabla_{fX'} Y)_p = f(p)(\nabla_{X'} Y)_p = (\nabla_{X'} Y)_p$$

This shows that $\nabla_X Y = \nabla_{X'} Y$ on U. Next, note that $fY = fY'$ on M implies that $\nabla_X(fY) = \nabla_X(fY')$, so the Leibniz rule and the facts that $f(p) = 1$, $X_p(f) = 0$ imply that $(\nabla_X Y)_p = (\nabla_X Y')_p$. Since p was taken to be an arbitrary point in U, $\nabla_X Y = \nabla_X Y'$ on U, and this completes the check of the claim.

2.2.1 Remark In a moment, we will refine the above discussion and show that, for a given point $p \in M$, the value of $(\nabla_X Y)_p$ depends only on X_p and the restriction of Y along a smooth curve $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ and $\gamma'(0) = X_p$. Indeed, this is a consequence of the expression of the connection (2.2.4).

Choose a chart $(U, \varphi = (x^1, \ldots, x^n))$ of M around p. We know from the above that $\nabla_X Y|_U = \nabla_{X|_U}(Y|_U)$. Write

$$X|_U = \sum_j a^j \frac{\partial}{\partial x^j} \quad \text{and} \quad Y|_U = \sum_k b^k \frac{\partial}{\partial x^k}$$

for a^j, $b^k \in C^\infty(U)$. Then, using the defining properties of a connection, in the open set U,

$$\begin{aligned}
\nabla_X Y &= \nabla_X \left(\sum_k b^k \frac{\partial}{\partial x^k} \right) \\
&= \sum_k X(b^k) \frac{\partial}{\partial x^k} + b^k \nabla_X \frac{\partial}{\partial x^k} \\
&= \sum_{j,k} a^j \frac{\partial b^k}{\partial x^j} \frac{\partial}{\partial x^k} + \sum_{j,k} a^j b^k \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} \\
&= \sum_{i,j} a^j \frac{\partial b^i}{\partial x^j} \frac{\partial}{\partial x^i} + \sum_{i,j,k} a^j b^k \Gamma^i_{jk} \frac{\partial}{\partial x^i},
\end{aligned}$$

where we have set

$$\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} = \sum_i \Gamma^i_{jk} \frac{\partial}{\partial x^i}.$$

It follows that the local representation of $\nabla_X Y$ in the chart (U, φ) is

(2.2.2)
$$\nabla_X Y = \sum_i \left(\sum_j a^j \frac{\partial b^i}{\partial x^j} + \sum_{j,k} \Gamma^i_{jk} a^j b^k \right) \frac{\partial}{\partial x^i}.$$

In particular,

(2.2.3)
$$(\nabla_X Y)_p = \sum_i \left(\sum_j a^j(p) \frac{\partial b^i}{\partial x^j}(p) + \sum_{j,k} \Gamma^i_{jk}(p) a^j(p) b^k(p) \right) \frac{\partial}{\partial x^i}\bigg|_p.$$

It is also convenient to rewrite the preceding formula in the following form

(2.2.4)
$$(\nabla_X Y)_p = \sum_i \left(X_p(b^i) + \sum_{j,k} \Gamma^i_{jk}(p) a^j(p) b^k(p) \right) \frac{\partial}{\partial x^i}\bigg|_p.$$

Note that this formula involves only the values of the a^j, b^k at p, and the directional derivatives of the b^i in the direction of X_p, so the claim in Remark 2.2.1 is checked.

The smooth functions Γ^i_{jk} are called the *Christoffel symbols* of ∇ with respect to the chosen chart. The Christoffel symbols of a connection satisfy a complicated transformation rule upon change of coordinates, which will be used in the proof of Proposition 2.3.1. For the moment, we just want to remark that the Christoffel symbols can be used to specify a connection locally. For instance, one could set Γ^i_{jk} identically zero in a given chart (U, φ) and then define a connection for vector fields on U. Doing this for a family of charts whose domains cover the manifold, and noting that a convex linear combination of connections is still a connection, a smooth partition of unity can be thus used to define a global connection in M in analogy with the argument in tbe proof of Proposition 1.2.3. This proves that connections exist in any given manifold.

Rather than insisting on the argument of the preceding paragraph, it is better to use Proposition 2.2.5 below in order to construct a connection in a given manifold. Indeed, in an n-dimensional smooth manifold, we need n^3 smooth functions Γ^i_{jk} to specify a connection locally, and we need n^2 smooth functions g_{ij} to specify a Riemannian metric locally, recall (1.2.1). Even taking into account equivalence classes of such objects, it is apparent that there exist "more" connections in a given smooth manifold than the already large amount of available Riemannian metrics. The point is that, as shown by the next proposition, a Riemannian manifold admits a preferred connection.

2.2.5 Proposition *Let (M, g) be a Riemannian manifold. Then there exists a unique connection ∇ in M, called the* Levi-Cività connection*, such that:*
a. $Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$*, and*
b. $\nabla_X Y - \nabla_Y X - [X, Y] = 0$
for all vector fields X, Y, $Z \in \Gamma(TM)$.

Proof. The strategy of the proof is to first use the two conditions in the statement to deduce a formula for ∇. This formula is called the *Koszul formula*, and this proves uniqueness. The next steps, which are easy but tedious and will be skipped, are to use the Koszul formula to define the connection, and to check that the defined object indeed satisfies the defining conditions of a connection and the conditions in the statement of this theorem.

Let X, Y and Z be vector fields in M. The so-called permutation trick is to use condition (a) to write

$$\begin{aligned} Xg(Y, Z) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ Yg(Z, X) &= g(\nabla_Y Z, X) + g(Z, \nabla_Y X) \\ -Zg(X, Y) &= -g(\nabla_Z X, Y) - g(X, \nabla_Z Y), \end{aligned}$$

add up these equations, and use condition (b) to arrive at the Koszul formula:

$$\begin{aligned} &g(\nabla_X Y, Z) = \\ (2.2.6) \qquad &\frac{1}{2}\big(Xg(Y, Z) + Yg(Z, X) - Zg(X, Y) + g([X, Y], Z) + g([Z, X], Y) + g([Z, Y], X)\big) \end{aligned}$$

Note that this formula uniquely defines $\nabla_X Y$, since Z is arbitrary and g is nondegenerate. □

The condition (a) in Proposition 2.2.5 is usually refered to as saying that the connection ∇ is *compatible with the metric* g, or that ∇ is a *metric connection.* The condition (b) expresses the fact that the *torsion* of ∇, which is defined as the left-hand side therein, is null.

Henceforth, in this book, for a given Riemannian manifold, we will always use the Levi-Cività connection in order to differentiate vector fields.

2.2.7 Example Consider the upper half-plane $\mathbf{R}^2_+ = \{\,(x,y)\in\mathbf{R}^2 \mid y>0\,\}$ endowed with the Riemannian metric $g=\frac{1}{y^2}(dx^2+dy^2)$. In this example, we show a practical method to compute the Levi-Cività connection of $(\mathbf{R}^2_+,g)$. Start with $g(\frac{\partial}{\partial x},\frac{\partial}{\partial x})=\frac{1}{y^2}$: differentiate it with respect to y and use Proposition 2.2.5(a) to write

$$2g\left(\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial x},\frac{\partial}{\partial x}\right)=\frac{\partial}{\partial y}\left(\frac{1}{y^2}\right)=-2\frac{1}{y^3},$$

so

$$g\left(\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial x},\frac{\partial}{\partial x}\right)=-\frac{1}{y^3}; \tag{2.2.8}$$

similarly, differentiate it with respect to x to get

$$g\left(\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial x},\frac{\partial}{\partial x}\right)=0.$$

Next, consider $g(\frac{\partial}{\partial y},\frac{\partial}{\partial y})=\frac{1}{y^2}$; differentiation with respect to x and y yields respectively

$$g\left(\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial y},\frac{\partial}{\partial y}\right)=0,\qquad g\left(\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial y},\frac{\partial}{\partial y}\right)=-\frac{1}{y^3}. \tag{2.2.9}$$

We use Proposition 2.2.5(b) in the form of

$$\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial y}-\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial x}=\left[\frac{\partial}{\partial x},\frac{\partial}{\partial y}\right]=0,$$

where the last equality holds because $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ are coordinate vector fields. Now differentiation of $g(\frac{\partial}{\partial x},\frac{\partial}{\partial y})=0$ gives that

$$g\left(\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial x},\frac{\partial}{\partial y}\right)=-g\left(\frac{\partial}{\partial x},\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial y}\right)=-g\left(\frac{\partial}{\partial x},\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial x}\right)=\frac{1}{y^3},$$

where we have used (2.2.8) in the last equality, and it also gives

$$g\left(\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial y},\frac{\partial}{\partial x}\right)=-g\left(\frac{\partial}{\partial y},\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial x}\right)=0,$$

where we have used the first formula of (2.2.9) in the last equality. Since $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ are orthogonal everywhere, it easily follows from the above formulas that

$$\left\{\begin{array}{lcrr}\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial x} & = & & \frac{1}{y}\frac{\partial}{\partial y}\\ \nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial y} & = & -\frac{1}{y}\frac{\partial}{\partial x} & \\ \nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial y} & = & & -\frac{1}{y}\frac{\partial}{\partial y}\end{array}\right.$$

⋆

2.3 Parallel transport along a curve

Let (M,g) be a Riemannian manifold, and denote by ∇ its Levi-Cività connection.

A *vector field along a curve* $\gamma : I \to M$, $I \subset \mathbf{R}$ an interval, is a map $X : I \to TM$ such that $X(t) \in T_{\gamma(t)}M$ for all t. If γ is a smooth curve, the most obvious example of a vector field along γ is its tangent vector field $\gamma'(t)$. In general, if γ is a proper embedding, then any vector field along γ can be extended to a smooth vector field in M defined on a neighborhood of the image of γ. On the other hand, if γ is not a proper embedding, then there are vector fields along γ that do not come from vector fields defined on open subsets of M. An example is given by taking γ to be the inclusion $(0,+\infty) \to \mathbf{R}$ and $X(t) = \frac{1}{t}\frac{d}{dt}$.

The set of smooth vector fields along a curve $\gamma : I \to M$ will be denoted $\Gamma(\gamma^*TM)$. The connection ∇ in M induces a derivative of vector fields along γ as follows.

2.3.1 Proposition *Let $\gamma : I \to M$ be a smooth curve. Then there exists a unique linear map $\frac{\nabla}{dt} : \Gamma(\gamma^*TM) \to \Gamma(\gamma^*TM)$, called the* covariant derivative along γ, *satisfying the following conditions:*

a. $\frac{\nabla}{dt}(fX) = \frac{df}{dt}X + f\frac{\nabla}{dt}X$ for every smooth function $f : I \to \mathbf{R}$.

b. If X admits an extension to a vector field $\bar{X}$ defined on an open subset U of M, then

$$\left(\frac{\nabla}{dt}X\right)(t) = (\nabla_{\gamma'(t)}\bar{X})_{\gamma(t)}$$

for every t satisfying $\gamma(t) \in U$.

Proof. We first prove the uniqueness result. Suppose first that the image of γ lies in the domain of one chart $(U, \varphi = (x^1,\ldots,x^n))$. Then we can write $\gamma(t) = (x^1(t),\ldots,x^n(t))$, so

$$\gamma'(t) = \sum_j (x^j)'(t)\frac{\partial}{\partial x^j}\Big|_{\gamma(t)}.$$

If X is a vector field along γ, we can also write

$$X(t) = \sum_k a^k(t)\frac{\partial}{\partial x^k}\Big|_{\gamma(t)}.$$

Note that, although in general X cannot be extended to a vector field defined on an open set of M, X is written as a linear combination of vector fields that admit such extensions. So, if we have a linear map as in the statement, then

$$\begin{aligned}
\frac{\nabla}{dt}X &= \sum_k (a^k)'\frac{\partial}{\partial x^k} + a^k\nabla_{\gamma'(t)}\frac{\partial}{\partial x^k} \\
&= \sum_i (a^i)'\frac{\partial}{\partial x^i} + \sum_{j,k} a^k (x^j)'\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^k} \\
&= \sum_i (a^i)'\frac{\partial}{\partial x^i} + \sum_{i,j,k} a^k (x^j)'\Gamma^i_{jk}\frac{\partial}{\partial x^i} \\
&= \sum_i \left((a^i)' + \sum_{j,k}\Gamma^i_{jk}(x^j)'a^k\right)\frac{\partial}{\partial x^i} \qquad (2.3.2)
\end{aligned}$$

In general, one sees by an argument analogous to that used in section 2.2 that $(\frac{\nabla}{dt}X)|_J$ depends only on $X|_J$ for any open subinterval J of I, and the image of γ can be covered by finitely many domains of charts, so the local expressions show that $\frac{\nabla}{dt}$ is uniquely defined, if it exists.

In order to prove existence, one uses the local expression to define $\frac{\nabla}{dt}$ in the domain of a local chart. Then, one easily checks that the defined map satisfies the two conditions in the statement. So far, we have existence on any open subset of M which is contained in the domain of a local chart, to which we can apply the uniqueness result from the first part of the proof. We finish by covering M with domains of local charts and noting that the locally defined covariant derivatives paste together to yield a globally defined object. □

A vector field X along a smooth curve $\gamma : I \to M$ is called *parallel* if $\frac{\nabla}{dt}X = 0$ on I. This definition can be obviously extended to include curves that are only piecewise smooth.

2.3.3 Proposition *Let $\gamma : I \to M$ be a piecewise smooth curve, and let $t_0 \in I$. Given a vector $v \in T_{\gamma(t_0)}M$, there exists a unique parallel vector field X along γ such that $X(t_0) = v$.*

Proof. Suppose first that I is compact. The image of γ can be covered by finitely many domains of charts of M. Thus, without loss of generality, we may assume that the image of γ lies in the domain of one chart $(U, \varphi = (x^1, \ldots, x^n))$. Write $\gamma(t) = (x^1(t), \ldots, x^n(t))$ and

$$X(t) = \sum_k a^k(t) \frac{\partial}{\partial x^k}\Big|_{\gamma(t)}.$$

Then, equation (2.3.2) implies that $\frac{\nabla}{dt}X = 0$ is equivalent to

$$(a^i)' + \sum_{j,k} \Gamma^i_{jk}(x^j)'a^k = 0 \tag{2.3.4}$$

for all i. This is a system of ordinary linear differential equations of first order in the unknowns $a^1, \ldots, a^n$, which is known to have unique solutions defined on all of I for given initial conditions. In our case, the initial conditions are given by $a_k(t_0) = dx^k(v)$.

In the general case, we can cover I by the union of a chain of increasing compact intervals, construct X along each compact interval, and use the uniqueness result to see that so constructed vector fields piece together to yield a global solution. □

It follows from the proof of the preceding proposition that the map that assigns to a vector $v \in T_{\gamma(t_0)}M$ a parallel vector field $X \in \Gamma(\gamma^*TM)$ with $X(t_0) = v$ is linear. Evaluating X at another time t_1 gives thus a linear map $P^\gamma_{t_1,t_0} : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M$ which will be called the *parallel translation map along γ from t_0 to t_1*.

2.3.5 Proposition *Let $\gamma : I \to M$ be a piecewise smooth curve. Then the parallel translation maps along γ enjoy the following properties:*

a. $P^\gamma_{t_0,t_0}$ is the identity map of $T_{\gamma(t_0)}M$;
b. $P^\gamma_{t_2,t_1} \circ P^\gamma_{t_1,t_0} = P^\gamma_{t_2,t_0}$ (chain rule);
c. $P^\gamma_{t_0,t_1} = (P^\gamma_{t_1,t_0})^{-1}$;
d. $P^\gamma_{t_1,t_0} : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M$ is an isometry;

for every t_0, t_1, $t_2 \in I$.

Proof. Assertions (a), (b) and (c) are immediate. We show that assertion (d) is a consequence of condition (a) in the definition of the Levi-Cività connection (in fact, it is equivalent to that condition) as follows. If X is a parallel vector field along γ, then $\frac{\nabla X}{dt} = 0$ along γ, so

$$\frac{d}{dt} g(X(t), X(t)) = 2g(\left(\frac{\nabla}{dt}X\right)(t), X(t)) = 0,$$

and the norm of X is constant along γ. □

2.3.6 Example We now use the result of Example 2.2.7 to describe the parallel transport map along the curve $\gamma(t) = (t, y_0)$ in $(\mathbf{R}^2_+, g)$, where $y_0 > 0$. Denote by $X(t) = a(t)\frac{\partial}{\partial x} + b(t)\frac{\partial}{\partial y}$ a smooth vector field along γ, where a, $b : \mathbf{R} \to \mathbf{R}$ are smooth functions. Then

$$\begin{aligned} \frac{\nabla}{dt}X &= a'\frac{\partial}{\partial x} + a\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial x} + b'\frac{\partial}{\partial y} + b\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial y} \\ &= \left(a' - \frac{b}{y_0}\right)\frac{\partial}{\partial x} + \left(b' + \frac{a}{y_0}\right)\frac{\partial}{\partial y}, \end{aligned}$$

so the condition that X be parallel is that

$$\begin{cases} a' = \omega b \\ b' = -\omega a \end{cases}$$

where $\omega = y_0^{-1}$. The general solution of this system of first-order ordinary differential equations is

$$\begin{aligned} a(t) &= a_0 \cos\omega t + b_0 \sin\omega t \\ b(t) &= -a_0 \sin\omega t + b_0 \cos\omega t \end{aligned}$$

where $(a(0), b(0)) = (a_0, b_0)$. It follows that

$$P^\gamma_{t,0}\left(a_0\frac{\partial}{\partial x} + b_0\frac{\partial}{\partial y}\right) = (a_0\cos\omega t + b_0\sin\omega t)\frac{\partial}{\partial x} + (-a_0\sin\omega t + b_0\cos\omega t)\frac{\partial}{\partial y}$$

which is merely rotation in the Euclidean sense at a constant rate; note that the rate $\omega \to \infty$ as $y_0 \to 0$. ⋆

2.4 Geodesics

Let (M, g) be a Riemannian manifold, and denote by ∇ its Levi-Cività connection.

A smooth curve $\gamma : I \to M$, $I \subset M$ an interval, is called a *geodesic* if and only if $\frac{\nabla}{dt}\gamma' = 0$ on I. Thus we require that the tangent vector field γ' be parallel along γ. According to 2.3.5(d), this implies that the length of γ' must be constant. We also refer to the latter property as saying that γ is a curve *parametrized with constant speed* or γ is a curve *parametrized proportionally to arc-length*. Observe that constant curves are geodesics.

We can get the local expression of the geodesic equation immediately from (2.3.4). Let $\gamma : I \to M$ be a smooth curve whose image lies in the domain of a chart $(U, \varphi = (x^1, \ldots, x^n))$ of M. Writing $\gamma(t) = (x^1(t), \ldots, x^n(t))$, we have that $\frac{\nabla}{dt}\gamma' = 0$ if and only if

$$(x^i)'' + \sum_{j,k}\Gamma^i_{jk}(x^j)'(x^k)' = 0 \qquad (2.4.1)$$

for all i. Note that this is a second order system of non-linear ordinary differential equations in the unknowns $x^1, \ldots, x^n$, for which we have a local existence and uniqueness result. Indeed, we quote the following theorem from [Spi70].

2.4.2 Theorem *Consider the second order system of ordinary differential equations*

$$\sigma'' = F\left(\sigma, \sigma'\right),$$

where $F : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$ is a smooth map, in the unknown $\sigma : I \to \mathbf{R}^n$, $I \subset \mathbf{R}$ an open interval. Then, given $(x_0, a_0) \in \mathbf{R}^n \times \mathbf{R}^n$, there exists a neighborhood $U \times V$ of (x_0, a_0) and $\delta > 0$

such that, for any $(x, a) \in U \times V$, there is a unique solution $\sigma_{x,a} : (-\delta, \delta) \to \mathbf{R}^n$ with initial conditions $\sigma_{x,a}(0) = x$ and $\sigma'_{x,a}(0) = a$. Moreover, the map $\Sigma : U \times V \times (-\delta, \delta) \to M$, defined by $\Sigma(x, a, t) = \sigma_{x,a}(t)$, is smooth.

It also follows from the theory of ordinary differential equations that any solution of the geodesic equation (2.4.1) is automatically smooth. Equation (2.4.1) has a particular homogeneity feature that we explore now. Namely, if $\gamma : (a, b) \to M$ is a solution of (2.4.1), then it is immediate to check that for every $k \in \mathbf{R} \setminus \{0\}$ the curve $\eta : (\frac{a}{k}, \frac{b}{k}) \to \mathbf{R}$ ($\eta : (\frac{b}{k}, \frac{a}{k}) \to \mathbf{R}$ if $k < 0$) defined by $\eta(t) = \gamma(kt)$ is also a solution.

2.4.3 Proposition *Given $p \in M$, there exists a neighborhood U of p and $\epsilon > 0$ such that, for any $q \in U$ and $v \in T_qM$ with $g_q(v, v)^{1/2} \leq \epsilon$, there is a unique geodesic $\gamma_v : (-2, 2) \to M$ such that $\gamma_v(0) = q$ and $\gamma'_v(0) = v$. Moreover, the map $\Gamma : \cup_{q\in U} B(0_q, \epsilon) \times (-2, 2) \to M$ defined by $\Gamma(v, t) = \gamma_v(t)$ is smooth.*

Proof. Let (V, φ) be a local chart of M around p, and consider the map $d\varphi : TM|_V \to \varphi(V) \times \mathbf{R}^n$. The geodesic equation in M corresponds via $d\varphi$ to a second order differential equation for curves on $\varphi(V) \times \mathbf{R}^n$, to which we apply Theorem 2.4.2. We deduce that there exists an open neighborhood W of 0_p in TM such that for every $v \in W$ there exists a unique geodesic $\gamma_v : (-\delta, \delta) \to M$ such that $\gamma_v(0) = \pi(v)$ and $\gamma'_v(0) = v$, where $\pi : TM \to M$ is the projection, and $\gamma_v(t)$ is smooth on $(v, t) \in W \times (-\delta, \delta)$. By continuity of g, we may shrink W and assume that it is of the form

$$W = \{ v \in TM|_U \ : \ g_{\pi(v)}(v, v)^{1/2} < \epsilon' \}$$

for some open neighborhood U of p in M and some $\epsilon' > 0$ (cf. Exercise 1). The homogeneity of the geodesic equation referred to above yields that multipliying the length of v by $\delta/2$ makes the interval of definition of γ_v to be multiplied by $2/\delta$. Therefore we can take $\epsilon = \epsilon'\delta/2$ and we are done. □

Henceforth, in this book, for $p \in M$ and $v \in T_pM$, we will denote by γ_v the unique geodesic with initial conditions $\gamma_v(0) = p$ and $\gamma'_v(0) = v$. Note that the homogeneity of the geodesic equation yields that $\gamma_{kv}(t) = \gamma_v(kt)$. It follows from Proposition 2.4.3 that there exists an open neighborhood Ω of the zero section in TM consisting of vectors v such that $\gamma_v(1)$ is defined. The *exponential map*

$$\exp : \Omega \to M$$

is defined by setting $\exp(v) = \gamma_v(1)$. It follows from the last assertion in Proposition 2.4.3 that the exponential map is smooth. Sometimes we will also write $\exp_p = \exp|_{T_pM}$ for $p \in M$. Now $\gamma_v(t) = \gamma_{tv}(1) = \exp_p(tv)$ for $v \in T_pM$ and sufficiently small t.

2.4.4 Proposition *Let $p \in M$. Then:*

a. *The exponential map exp_p maps an open neighborhood of $0_p \in T_pM$ diffeomorphically onto an open neighborhood of p in M.*
b. *There exists an open neighborhood U of p and $\epsilon > 0$ such that, for any $q \in U$, there exists a unique $v \in T_pM$ with $g_p(v, v)^{1/2} < \epsilon$ such that $\exp_p v = q$.*

Proof. We compute the differential $d(\exp_p)_{0_p} : T_{0_p}(T_pM) \to T_pM$. Recall that $\exp_p(tv) = \gamma_{tv}(1) = \gamma_v(t)$ for $v \in T_pM$. Differentiating this equation with respect to t at $t = 0$ yields that

$$d(\exp_p)_{0_p}(v) = \gamma'_v(0) = v. \tag{2.4.5}$$

Hence $d(\exp_p)_{0_p}$ is the identity, where as usual we have identified $T_{0_p}(T_pM)$ with T_pM. It follows from the inverse function theorem that $\exp_p$ maps an open neighborhood of 0_p in T_pM, which can be taken of the form $B(0_p, \epsilon)$ for some $\epsilon > 0$, diffeomorphically onto an open neighborhood of p in M. Parts (a) and (b) follow. □

The neighborhood of p given in the previous proposition is usually called a *normal neighborhood* of p. Hence we have that any point in a normal neighborhood of p can be joined to p by a unique geodesic in that neighborhood. Next, we want to improve this result in the sense of connecting two movable points in a neighborhood of p by a geodesic. We need a lemma.

2.4.6 Lemma *Let $\pi : TM \to M$ be the projection. Then, given $p \in M$, the map*

$$\Phi : \Omega \to M \times M, \qquad \Phi(v) = (\pi(v), \exp(v))$$

is a local diffeomorphism from an open neighborhood W of 0_p in TM onto an open neighborhood of (p, p) in $M \times M$.

Proof. The result follows from the inverse function theorem if we can show that $d\Phi_{0_p} : T_{0_p}(TM) \to T_pM \oplus T_pM$ is an isomorphism. Each vector in the tangent space $T_{0_p}(TM)$ is the tangent vector at $t = 0$ to a curve c in TM passing through 0_p at $t = 0$. First, let $c(t) = tv \in TM$ where $v \in T_pM$. Then $d\Phi_{0_p}(c'(0)) = \frac{d}{dt}\big|_{t=0}\Phi(c(t)) = \frac{d}{dt}\big|_{t=0}(p, \exp_p(tv)) = (0, v)$ by equation (2.4.5). Next, let $c(t) = 0_{\gamma(t)} \in T_{\gamma(t)}M \subset TM$, where γ is a curve in M with $\gamma(0) = p$ and $\gamma'(0) = v \in T_pM$. Then $d\Phi_{0_p}(c'(0)) = \frac{d}{dt}\big|_{t=0}\Phi(0_{\gamma(t)}) = \frac{d}{dt}\big|_{t=0}(\gamma(t), \gamma(t)) = (v, v)$. The two calculations together imply that $d\Phi_{0_p}$ is surjective and hence, by dimensional reasons, an isomorphism. □

2.4.7 Proposition *Given $p \in M$, there exists an open neighborhood U of p and $\epsilon > 0$ such that:*

a. *For any x, $y \in U$, there exists a unique $v \in T_xM$ with $g_x(v, v)^{1/2} < \epsilon$ such that $\exp_x v = y$. Set $\gamma_v(t) = \exp_x(tv)$.*
b. *The map $\Psi : U \times U \times [0, 1] \to M$ defined by $\Psi(x, y, t) = \gamma_v(t)$ is smooth.*
c. *For all $x \in U$, the map $\exp_x$ is a diffeomorphism from $B(0_x, \epsilon)$ onto a normal neighborhood of x containing U.*

Proof. (a) Let W be a neighborhood of 0_p in TM such that the map of Lemma 2.4.6 is a diffeomorphism of W onto a neighborhood of (p, p) in $M \times M$. By shrinking W, if necessary, we may assume that $W = \cup_{x\in V} B(0_x, \epsilon)$ for some open neighborhood V of p and some $\epsilon > 0$. Let U be a neighborhood of p in M contained in V and such that $U \times U \subset \Phi(W)$. Then, for any $(x, y) \in U \times U$, there is a unique $v \in W$ such that $\Phi(v) = (x, y)$, meaning that there is a unique $v \in B(0_x, \epsilon)$ such that $\exp_x v = y$.

(b) This follows immediately from the fact that $\Psi(x, y, t) = \exp(t\Phi^{-1}(x, y))$.

(c) Since $B(0_x, \epsilon) \subset W$, the map Φ is a diffeomorphism from $B(0_x, \epsilon)$ onto its image. But, for fixed $x \in U$, $\Phi(v) = (x, \exp_x(v))$ for $v \in B(0_x, \epsilon)$. □

The set U in the preceding proposition is a normal neighborhood of each of its points; we will call such a set U an *ϵ-totally normal neighborhood* of p. Note that it is not claimed that the geodesic γ_v in the proposition is entirely contained in U. In section 6.6, we will prove that we can always find a possibly smaller totally normal neighborhood of p with that property.

2.4.8 Example In order to complete our analysis of the Riemannian manifold $(\mathbf{R}^2_+, g)$ of Examples 2.2.7 and 2.3.6, we now determine its geodesics. So let $\gamma(t) = (x(t), y(t))$ be a smooth curve in $\mathbf{R}^2_+$. Then $\gamma' = x'\frac{\partial}{\partial x} + y'\frac{\partial}{\partial y}$ and

$$\frac{\nabla}{dt}\gamma' = x''\frac{\partial}{\partial x} + x'\frac{\nabla}{dt}\frac{\partial}{\partial x} + y''\frac{\partial}{\partial y} + x'\frac{\nabla}{dt}\frac{\partial}{\partial y}.$$

We also have

$$\frac{\nabla}{dt}\frac{\partial}{\partial x} = x'\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial x} + y'\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial x} = -\frac{y'}{y}\frac{\partial}{\partial x} + \frac{x'}{y}\frac{\partial}{\partial y},$$

and

$$\frac{\nabla}{dt}\frac{\partial}{\partial y} = x'\nabla_{\frac{\partial}{\partial x}}\frac{\partial}{\partial y} + y'\nabla_{\frac{\partial}{\partial y}}\frac{\partial}{\partial y} = -\frac{x'}{y}\frac{\partial}{\partial x} - \frac{y'}{y}\frac{\partial}{\partial y},$$

so

$$\frac{\nabla}{dt}\gamma' = \Big(x'' - 2\frac{x'y'}{y}\Big)\frac{\partial}{\partial x} + \Big(y'' + \frac{x'^2 - y'^2}{y}\Big)\frac{\partial}{\partial y}.$$

Therefore the geodesic equations are

$$\left\{\begin{array}{l} x'' - 2\frac{x'y'}{y} = 0 \\ y'' + \frac{x'^2 - y'^2}{y} = 0 \end{array}\right. \tag{2.4.9}$$

Note that $x(t) = x_0$ is a solution of (2.4.9); indeed, the second equation gives that

$$\left(\frac{y'}{y}\right)' = \frac{y''y - y'^2}{y^2} = 0,$$

so $y(t) = y_0 e^{kt}$ where $y_0 > 0$ and $k \in \mathbf{R}$. This shows that the vertical lines are geodesics. Note that in the parametrization that we obtained, they are defined on all of $\mathbf{R}$.

Next, suppose that γ is a geodesic which is not a vertical line. By the uniqueness result for geodesics, it follows that $x'(t) \neq 0$ for all t in the domain of γ. The first equation of (2.4.9) then gives

$$\frac{x''}{x'} = 2\frac{y'}{y}$$

from where we get that

$$(\log|x'|)' = (2\log y)'$$

and hence that

$$x' = cy^2 \tag{2.4.10}$$

for some real constant c which may be assumed to be positive by reversing the orientation of γ, if necessary. Of course γ is parametrized with constant speed, which for simplicity we assume is 1; then $\frac{1}{y^2}(x'^2 + y'^2) = 1$; substituing (2.4.10) gives that

$$\frac{dy}{y\sqrt{1 - c^2y^2}} = \pm dt$$

Direct integration then yields

$$\operatorname{arcsech}(cy) = \pm t - t_0,$$

and changing the initial point we may assume that $t_0 = 0$. Then

$$y(t) = R\operatorname{sech} t \tag{2.4.11}$$

where $R = c^{-1} > 0$. Finally, equation (2.4.10) implies that

$$x(t) = x_0 + R\tanh t \tag{2.4.12}$$

for some $x_0 \in \mathbf{R}$. Note that equations (2.4.12) and (2.4.11) are defined on all of $\mathbf{R}$, and they parametrize the semi-circle of center $(x_0, 0)$ and radius R in $\mathbf{R}^2_+$.

Any geodesic of $(\mathbf{R}_2^+, g)$ is of one of the above types. Indeed, given initial conditions for a geodesic, it is readily seen that there exists a (unique) vertical line or semi-circle as above satisfying the given initial conditions. ⋆

2.5 Isometries and Killing fields

It is more or less clear that isometries should preserve any object canonically associated to a Riemannian manifold. Let (M, g) and (M', g') be Riemannian manifolds, denote by ∇ and ∇' the corresponding Levi-Cività connections, and let $f : M \to M'$ be an isometry. It follows from the Koszul formula (2.2.6) that f maps ∇ to ∇' in the sense that

$$\nabla'_{f_*X} f_*Y = f_*(\nabla_X Y)$$

where X, $Y \in \Gamma(TM)$. In particular, if $\gamma : I \to M$ is a geodesic of (M, g) then $f \circ \gamma : I \to M'$ is a geodesic of (M', g').

It is interesting to rephrase the last assertion in terms the exponential map. Namely, if f is an isometry of (M, g), $p \in M$ and $v \in T_pM$ lies in the domain of $\exp_p$, then $df_p(v)$ lies in the domain of $\exp_{f(p)}$ and

$$f(\exp_p(v)) = \exp_{f(p)}(df_p(v)).$$

In particular, if p is a fixed point of f then, on a normal neighborhood of p, we can write

$$f = \exp_p \circ\, df_p \circ \exp_p^{-1};$$

namely, $\exp_p^{-1}$ defines a local chart on a normal neighborhood of p that linearizes f.

A *Killing vector field* (sometimes, simply a *Killing field*) on a Riemannian manifold (M, g) is a smooth vector field X on M whose local flow $\{\varphi_t\}$ consists of local isometries of M, namely, $\varphi_t^* g = g$ wherever defined. By differentiation with respect to t, we immediately see that this condition is equivalent to the vanishing of Lie derivative of g with respect to X,

$$L_X g = 0,$$

or equivalently,

$$Xg(Y, Z) = g([X, Y], Z) + g(Y, [X, Z]) \qquad (2.5.1)$$

for every Y, $Z \in \Gamma(TM)$.

2.5.2 Proposition *Let (M, g) be a Riemannian manifold.*

a. The set of Killing fields on M form a Lie subalgebra of the Lie algebra of smooth vector fields on M.

b. A smooth vector field $X \in \Gamma(TM)$ is a Killing field if and only if

$$g(\nabla_Y X, Z) + g(\nabla_Z X, Y) = 0$$

for every Y, $Z \in \Gamma(TM)$, i. e. $(\nabla X)_p$ is skew-symmetric as a linear operator on T_pM for all $p \in M$.

Proof. (a) The set of Killing fields on M is a subspace of $\Gamma(TM)$ because $L_X g = 0$ is linear in X, and closed under the Lie bracket because $L_{[X,Y]} = [L_X, L_Y]$ for all X, $Y \in \Gamma(TM)$.

(b) Using that the Levi-Cività connection is compatible with the metric and has no torsion (Proposition 2.2.5(a) and (b)), equation (2.5.1) is seen to be equivalent to

$$g(\nabla_X Y, Z) + g(Y, \nabla_X Z) = g(\nabla_X Y - \nabla_Y X, Z) + g(Y, \nabla_X Z - \nabla_Z X),$$

which implies the result. □

Recall that the set $\mathrm{Isom}(M,g)$ of all isometries of a Riemannian manifold (M,g) forms a subgroup of the group of all diffeomorphisms of M, which has the structure of a Lie group with respect to the compact-open topology; moreover, the map $\mathrm{Isom}(M,g)\times M\to M$ is smooth [KN96]. In particular, if all Killing fields are complete, then the Lie algebra of $\mathrm{Isom}(M,g)$ is naturally identified with the Lie algebra of Killing fields of M.

2.5.3 Remark In chapter 3 we will see that Killing fields are complete if M is e. g. compact. It follows from exercise 6 of chapter 5 that the dimension of the Lie algebra of Killing fields on M is bounded by $\frac{1}{2}n(n+1)$, where $n=\dim M$.

2.6 Induced connections

At this juncture, it is convenient to introduce the following extension of Proposition 2.3.1. We will be using it especially in the case $\dim N = 2$.

2.6.1 Proposition *Let N be a smooth manifold, and let $\varphi : N \to M$ be a smooth map. Then there exists a unique bilinear map $\nabla^\varphi : \Gamma(TN)\times\Gamma(\varphi^*TM)\to\Gamma(\varphi^*TM)$, called the* induced connection along *φ, satisfying the following conditions:*
a. $\nabla^\varphi_{fX}Y = f\nabla^\varphi_X Y$;
b. $\nabla^\varphi(fY) = X(f)Y + f\nabla^\varphi_X Y$;
c. If Y admits an extension to a vector field $\hat{Y}$ defined on an open subset U of M, then

$$\left(\nabla^\varphi_X Y\right)_p = \left(\nabla_{d\varphi(X_p)}\hat{Y}\right)_{\varphi(p)}$$

for every $p\in\varphi^{-1}(U)$;
where $X\in\Gamma(TN)$, $Y\in\Gamma(\varphi^*TM)$ and $f : N\to\mathbf{R}$ is a smooth function.

Proof. Similar to the proof of Proposition 2.3.1. □

2.6.2 Proposition *Let $\varphi : N\to M$ be a smooth map, let X, $Y\in\Gamma(TN)$ be vector fields in N and let U, $V\in\Gamma(\varphi^*TM)$ be vector fiels along φ. Then the following identities hold:*

$$\nabla^\varphi_X(\varphi_*Y) - \nabla^\varphi_Y(\varphi_*X) - \varphi_*[X,Y] = 0, \quad \textit{and}$$
$$X\,g(U,V) - g(\nabla^\varphi_X U, V) - g(\nabla^\varphi_X V, U) = 0.$$

Proof. One checks that the left-hand sides of both expressions are $C^\infty(M)$-linear in the arguments, so it suffices to check the formulae in the case in which X, Y are coordinate vector fields on N and U, V are restrictions of coordinate vector fields of M along φ. In this special case, the desired results reduce to known properties of the Levi-Civitá connection on M. □

2.7 Connections on vector bundles ★

Let $E\to M$ be a vector bundle over a smooth manifold M. A *(linear) connection* on E is an $\mathbf{R}$-bilinear map $\nabla : \Gamma(TM)\times\Gamma(E)\to\Gamma(E)$ such that
a. $\nabla_{fX}\xi = f\nabla_X\xi$, and
b. $\nabla_X(f\xi) = X(f)\xi + f\nabla_X\xi$ (Leibniz rule)

for every $X \in \Gamma(TM)$, $\xi \in \Gamma(E)$ and $f \in C^\infty(M)$. If, in addition, E is a Riemannian vector bundle, i.e. it is endowed with an inner product structure $\langle\cdot,\cdot,\rangle$, then a connection ∇ on E is called *compatible with the metric* g if

$$X\langle\xi,\eta\rangle = \langle\nabla_X\xi,\eta\rangle + \langle\xi,\nabla_X\eta\rangle$$

for all $X \in \Gamma(TM)$, ξ, $\eta \in \Gamma(E)$.

Similarly to the special case $E = TM$, an arbitrary connection on $E \to M$ has the property that $(\nabla_X\xi)_p$ depends only on the value of X at p and the values of ξ along a given smooth curve in M which is tangent to X_p at some point.

Finally, if ∇ is a connection on $E \to M$ and $\varphi : N \to M$ is a smooth map, then there is an induced connection along φ, denoted ∇^φ, on the induced bundle $\varphi^*E \to N$, in analogy with the result in Proposition 2.6.1.

2.8 Examples

The Euclidean space

We claim that the Levi-Cività connection ∇ in $\mathbf{R}^n$ coincides with the usual derivative. In fact, let $(x^1,\ldots,x^n)$ denote the standard global coordinates in $\mathbf{R}^n$. We have that

$$g\left(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}\right) = \delta_{ij} \quad \text{and} \quad \left[\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}\right] = 0$$

for all i, j. Plugging these relations into the Koszul formula (2.2.6) gives that $\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} = 0$ for all i, j, namely, all the Christoffel symbols $\Gamma^i_{jk} = 0$. If

$$X = \sum_j a^j\frac{\partial}{\partial x^j} \quad \text{and} \quad Y = \sum_k b^k\frac{\partial}{\partial x^k},$$

for a^j, $b^k \in C^\infty(\mathbf{R}^n)$, then, using formula (2.2.2),

$$\nabla_X Y = \sum_i\left(\sum_j a^j\frac{\partial b^i}{\partial x^j}\right)\frac{\partial}{\partial x^i} = X(Y) = dY(X),$$

proving the claim. We also get, from equation (2.3.4), that a vector field X along a curve $\gamma : [a,b] \to M$, given as

$$X(t) = \sum_k a^k(t)\frac{\partial}{\partial x^k}\Big|_{\gamma(t)},$$

is parallel if and only the a_k are constant functions, namely, the parallel vector fields in $\mathbf{R}^n$ are the constant vector fields. It follows that the parallel transport map along γ from a to b is given by the differential of the translation map, that is,

$$P^\gamma_{b,a} = d(\tau_v)_{\gamma(a)},$$

where τ_v is the translation in $\mathbf{R}^n$ by the vector $v = \gamma(b) - \gamma(a)$, and, in particular, is independent of the curve γ joining $\gamma(a)$ and $\gamma(b)$. Finally, the geodesic equation (2.4.1) in $\mathbf{R}^n$ is

$$(x^i)'' = 0$$

for all i, so the geodesics are the lines. Hence

$$\exp_p(v) = p + v$$

for $p \in \mathbf{R}^n$ and $v \in T_p\mathbf{R}^n = \mathbf{R}^n$.

Product Riemannian manifolds

Let (M_i, g_i), where $i = 1, 2$, denote two Riemannian manifols and consider the product Riemannian manifold $(M, g) = (M_1, g_1) \times (M_2, g_2)$. Let $U_i \in \Gamma(TM_i)$, where $i = 1, 2$, be arbitrary vector fields. Of course, U_1 and U_2 can be identified with vector fields on M, and it follows from the construction of (M, g) that $[U_1, U_2] = 0$ and $g(U_1, U_2) = 0$ in M.

Now, suppose that X, Y, $Z \in \Gamma(TM)$ can be decomposed as $X = X_1 + X_2$, $Y = Y_1 + Y_2$ and $Z = Z_1 + Z_2$, where X_i, Y_i, $Z_i \in \Gamma(TM_i)$ for $i = 1, 2$ (not every vector field on M admits such a decomposition!). Note that

$$Xg(Y, Z) = X_1 g_1(Y_1, Z_1) + X_2 g_2(Y_2, Z_2)$$

and

$$g([X, Y], Z) = g_1([X_1, Y_1], Z_1) + g_2([X_2, Y_2], Z_2).$$

It then follows from the Koszul formula (2.2.6) applied three times that

$$\begin{aligned} g(\nabla_X Y, Z) &= g_1(\nabla^1_{X_1} Y_1, Z_1) + g_2(\nabla^2_{X_2} Y_2, Z_2) \\ &= g(\nabla^1_{X_1} Y_1 + \nabla^2_{X_2} Y_2, Z), \end{aligned}$$

where ∇ denotes the Levi-Cività connection of M and ∇^i denotes the Levi-Cività connection of M_i for $i = 1, 2$. Since g is nondegenerate and any tangent vector to M can be extended to a vector field Z which decomposes as $Z_1 + Z_2$, this calculation yields the following formula for the Levi-Cività connection of a Riemannian product:

$$\nabla_X Y = \nabla^1_{X_1} Y_1 + \nabla^2_{X_2} Y_2. \tag{2.8.1}$$

It follows from this formula that the Christoffel symbol Γ^i_{jk} of ∇ is zero unless all the three indices i, j, k correspond to coordinates of the same factor M_ℓ, where $\ell = 1$ or 2, in which case Γ^i_{jk} is a function on M_ℓ and a Christoffel symbol of ∇^ℓ. Therefore if γ is a curve in M with components γ_1 in M_1 and γ_2 in M_2, and X is a vector field along γ, then we can decompose $X = X_1 + X_2$ where X_i is a vector field along γ_i, and equation (2.3.2) gives $\frac{\nabla X}{dt} = \frac{\nabla X_1}{dt} + \frac{\nabla X_2}{dt}$. In particular, X is parallel along γ if and only if X_i is parallel along M_i for $i = 1, 2$. As $\gamma'(t) = \gamma_1'(t) + \gamma_2'(t)$, in particular yet, γ is a geodesic if and only if γ_i is a geodesic of M_i for $i = 1, 2$.

Riemannian submanifolds and isometric immersions

Let (M, g), $(\overline{M}, \overline{g})$ be Riemannian manifolds, and suppose that $\iota : M \to \overline{M}$ is an isometric immersion. We would like to relate the Levi-Cività connections ∇ of M and $\overline{\nabla}$ of $\overline{M}$. Since this is a local problem, we can work in a neighborhood of a point $p \in M$ and assume that ι is the inclusion map. Now the tangent bundle TM is a subbundle of $T\overline{M}|_M$, the metric g is the restriction of $\overline{g}$, and every vector field on M admits an extension to a vector field on $\overline{M}$.

Let X, Y and Z be vector fields on M, and let $\overline{X}$, $\overline{Y}$ and $\overline{Z}$ be extensions of those vector fields to vector fields on $\overline{M}$. Note that $[\overline{X}, \overline{Y}]$ is an extension of $[X, Y]$ to a vector field on $\overline{M}$. It follows from two applications of the Koszul formula (2.2.6) that

$$\begin{aligned} 2\overline{g}((\nabla_X Y)_p, Z_p) &= 2g((\nabla_X Y)_p, Z_p) \\ &= \mathfrak{S} \ \pm X_p\, g(Y, Z) \pm g([X, Y]_p, Z_p) \\ &= \mathfrak{S} \ \pm \overline{X}_p\, \overline{g}(\overline{Y}, \overline{Z}) \pm \overline{g}([\overline{X}, \overline{Y}]_p, \overline{Z}_p) \\ &= 2\overline{g}((\overline{\nabla}_{\overline{X}} \overline{Y})_p, \overline{Z}_p) \\ &= 2\overline{g}((\overline{\nabla}_{\overline{X}} \overline{Y})_p, Z_p), \end{aligned}$$

where $\mathfrak{S}$ denotes cyclic summation in X, Y, Z. Since $(\nabla_X Y)_p \in T_pM$ and Z_p can be any element of T_pM, it follows that

$$(\nabla_X Y)_p = \Pi_p \left((\overline{\nabla}_{\overline{X}} \overline{Y})_p \right), \tag{2.8.2}$$

where $\Pi_p : T_p\overline{M} \to T_pM$ is the orthogonal projection.

The most important case is that of Riemannian submanifolds of Euclidean space. If M is a Riemannian submanifold of $\mathbf{R}^n$, then formula (2.8.2) implies that a smooth curve γ in M is a geodesic of M if and only if its second derivative γ'' in $\mathbf{R}^n$ is everywhere normal to M; in other words, the geodesics of M are the "curves with normal acceleration".

The sphere S^n

Let $p \in S^n$ and $v \in T_pS^n$. We now determine the unique geodesic γ of S^n with initial conditions $\gamma(0) = p$ and $\gamma'(0) = v$. If $v = 0$, then γ is a constant curve, so we may assume that $v \neq 0$. Since p and v are orthogonal vectors in $\mathbf{R}^{n+1}$, they span a 2-dimensional subspace which we denote by E. Let $f : \mathbf{R}^{n+1} \to \mathbf{R}^{n+1}$ be the linear reflection on E. Then f is an orthogonal transformation of $\mathbf{R}^{n+1}$ and leaves S^{n+1} invariant. Now every orthogonal transformation of $\mathbf{R}^{n+1}$ is an isometry. Since S^{n+1} has the induced metric from $\mathbf{R}^{n+1}$, f restricts to an isometry of S^n which we denote by the same letter. Owing to the fact that an isometry maps geodesics to geodesics, the curve $\tilde{\gamma} = f \circ \gamma$ is a geodesic of S^n. Since f leaves E pointwise fixed, the initial conditions of $\tilde{\gamma}$ are $\tilde{\gamma}(0) = f(\gamma(0)) = f(p) = p$ and $\tilde{\gamma}'(0) = f(\gamma'(0)) = f(v) = v$, namely, the same as those of γ. By the uniqueness of geodesics with given initial conditions, we have that $\tilde{\gamma} = \gamma$, or, what is the same, $f(\gamma(t)) = \gamma(t)$ for all t in the domain of γ. It follows that γ is contained in E and thus must coincide with the great circle $S^n \cap E$ parametrized with constant speed on its domain of definition. This argument shows that the great circles are locally geodesics; but then, the great circles are geodesics.

In particular, the geodesics of S^n parametrized by arc-length are periodic of period 2π. Finally, we have the formula

$$\exp_p(v) = \cos(||v||)p + \sin(||v||)\frac{v}{||v||}$$

for $v \neq 0$.

Riemannian coverings

Let $\pi : (\tilde{M}, \tilde{g}) \to (M, g)$ be a Riemannian covering.

2.8.3 Proposition *The geodesics of (M, g) are the projections of the geodesics of $(\tilde{M}, \tilde{g})$, and the geodesics of $(\tilde{M}, \tilde{g})$ are the liftings of the geodesics of (M, g).*

Proof. Suppose $\tilde{\gamma}$ and γ are smooth curves in $\tilde{M}$, M such that $\pi \circ \tilde{\gamma} = \gamma$. Since π is a local isometry, it maps a sufficiently small arc of $\tilde{\gamma}$ isometrically onto a small arc of γ. It follows that $\tilde{\gamma}$ is a geodesic if and only if γ is a geodesic. This shows that the classes of curves described in the statement of the proposition are indeed geodesics. Now we need only to remark that every smooth curve in M is the projection of any of its smooth liftings in $\tilde{M}$, and every smooth curve in $\tilde{M}$ is the smooth lifting of its projection to M. $\square$

The real projective space

We apply Proposition 2.8.3 to the Riemannian covering map $\pi : S^n \to \mathbf{R}P^n$. The geodesics of S^n have already been determined as being the great circles parametrized with constant speed, so the geodesics of $\mathbf{R}P^n$ are the projections of those. In particular, since π identifies antipodal points of S^n, the geodesics of $\mathbf{R}P^n$ paramerized by arc-length are periodic of period π.

Flat tori

Let Γ be a lattice in $\mathbf{R}^n$ and consider the induced Riemannian metric g_Γ on T^n. We apply Proposition 2.8.3 to the Riemannian covering map $\pi : \mathbf{R}^n \to (T^n, g_\Gamma)$ to deduce that the geodesics of (T^n, g_Γ) are simply the projections of the straight lines in $\mathbf{R}^n$. In this way, we see that some geodesics of (T^n, g_Γ) are periodic and some are dense in T^n.

Next, let Γ' be another lattice in $\mathbf{R}^n$. We have already remarked that (T^n, g_Γ) and $(T^n, g_{\Gamma'})$ are generally non-isometric. Nevertheless, there existsa linear transformation f of $\mathbf{R}^n$ that maps Γ to Γ', and hence induces a diffeomorphism $\bar f : \mathbf{R}^n/\Gamma \to \mathbf{R}^n/\Gamma'$ such that the diagram

$$\begin{array}{ccc} \mathbf{R}^n & \xrightarrow{f} & \mathbf{R}^n \\ \downarrow & & \downarrow \\ \mathbf{R}^n/\Gamma & \xrightarrow{\bar f} & \mathbf{R}^n/\Gamma' \end{array}$$

is commutative. In general, $\bar f$ is not an isometry, but since f maps straight lines to straight lines, $\bar f$ maps the geodesics of (T^n, g_Γ) to the geodesics of $(T^n, g_{\Gamma'})$. Hence we get an example of two non-isometric metrics on the same smooth manifold with the same geodesics.

Lie groups ★

Let G be a Lie group and denote its Lie algebra by $\mathfrak{g}$. In this example, we will describe the Levi-Cività connection associated to a bi-invariant metric on G. We start with a definition and a proposition.

We say that an inner product $\langle , \rangle$ on $\mathfrak{g}$ is ad-*invariant* if the identity

$$\langle \mathrm{ad}_Z X, Y \rangle + \langle X, \mathrm{ad}_Z Y \rangle = 0 \tag{2.8.4}$$

holds for every X, Y, $Z \in \mathfrak{g}$.

2.8.5 Proposition *Every* Ad-*invariant inner product on* $\mathfrak{g}$ *is* ad-*invariant, and the converse holds if* G *is connected.*

Proof. Let $\langle , \rangle$ be an inner product on $\mathfrak{g}$. It being Ad-invariant means that

$$\langle \mathrm{Ad}_g X, \mathrm{Ad}_g Y \rangle = \langle X, Y \rangle \tag{2.8.6}$$

for every $g \in G$ and X, $Y \in \mathfrak{g}$. In particular, taking $g = \exp tZ$ for $Z \in \mathfrak{g}$ and differentiating at $t = 0$ yields identity (2.8.4).

Assume now that G is connected and $\langle , \rangle$ is ad-invariant. Then (2.8.4) holds; note that what it is really saying is that $f'_{X,Y}(0) = 0$ for all X, $Y \in \mathfrak{g}$, where

$$f_{X,Y}(t) = \langle \mathrm{Ad}_{\exp tZ} X, \mathrm{Ad}_{\exp tZ} Y \rangle,$$

and from this information we will show that $f_{X,Y}(t) = f_{X,Y}(0)$. Indeed, since $t \mapsto \mathrm{Ad}_{\exp tZ}$ is a homomorphism,

$$f_{X,Y}(t+s) = f_{X',Y'}(t)$$

where $X' = \mathrm{Ad}_{\exp sZ}X$ and $Y' = \mathrm{Ad}_{\exp sZ}Y$. Differentiating this identity at $t = 0$ gives that $f'_{X,Y}(s) = f'_{X',Y'}(0) = 0$. Since $s \in \mathbf{R}$ is arbitrary, this implies that $f_{X,Y}$ is constant, as desired.

So far we have shown that (2.8.6) holds if g lies in the image of exp. But there exists an open neighborhood U of the identity of G contained in the image of exp, and it is known that U generates G as a group due to the connectedness of G. Since $g \mapsto \mathrm{Ad}_g$ is a homomorphism, this finally implies that (2.8.6) holds for every $g \in G$. □

Let g be a bi-invariant metric on G. Now we are ready to use the Koszul formula (2.2.6) to compute the Levi-Cività connection on left-invariant vector fields. Let X, Y, $Z \in \mathfrak{g}$. Since X and Y are left-invariant vector fields and g is a left-invariant metric, $g(X,Y)$ is a constant function on G. Therefore $Zg(X,Y) = 0$. Similarly, $Yg(Z,X) = Xg(Y,Z) = 0$. Regarding the other terms of (2.2.6), the preceding proposition shows that g_1 is an ad-invariant inner product on $\mathfrak{g}$, so

$$g([Z,X],Y) + g(X,[Z,Y]) = g_1(\mathrm{ad}_Z X, Y) + g_1(X, \mathrm{ad}_Z Y) = 0. \tag{2.8.7}$$

We deduce that

$$\nabla_X Y = \frac{1}{2}[X,Y] \tag{2.8.8}$$

for all X, $Y \in \mathfrak{g}$ (this formula shows in particular that $\nabla_X Y$ is also a left-invariant vector field, but this fact of course also follows from general properties of isometries, cf. section 2.5). An important application of this formula is that $\nabla_X X = 0$ for all $\in \mathfrak{g}$, and this means that every one-parameter subgroup of G through the identity is a geodesic. This is also equivalent to saying that the exponential map of G *qua* Lie group and the exponential map of G *qua* Riemannian manifold (G,g) coincide. It follows from the Hopf-Rinow theorem to be proved in the next chapter that the exponential map of a compact connected Lie group is surjective, see Theorem 3.3.2 and Corollary 3.3.6. Of course, the geodesics of G through an arbitrary point are left-translates of one-parameter subgroups, namely, of the form $t \mapsto g \exp tX$ for $g \in G$ and $X \in \mathfrak{g}$.

2.9 Exercises

1 Let (M,g) be a Riemannian manifold, consider its tangent bundle TM, and fix a point $p \in M$. Prove that any open neighborhood W of 0_p in TM contains a neighborhood of the form

$$\bigcup_{x\in U} B(0_x,\epsilon) = \{\, v \in TM|_U \;:\; g_{\pi(v)}(v,v)^{1/2} < \epsilon \,\}$$

for some open neighborhood U of p in M and some $\epsilon > 0$.

2 Let A, B be nowhere zero smooth functions on $\mathbf{R}^2$ and consider the Riemannian metric $g = A^2\,dx^2 + B^2\,dy^2$, where x, y are the standard coordinates on $\mathbf{R}^2$.
a. Compute the Christoffel symbols of g.
b. Write down the geodesic equations of g.

3 Let (x^i) be a system of local coordinates on a smooth manifold M which is equipped with a connection ∇, and consider the Christoffel symbols Γ^k_{ij} which are defined by $\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} = \sum_k \Gamma^k_{ij}\frac{\partial}{\partial x^k}$.

If $(\bar{x}^{i'})$ is another system of local coordinates on M, prove that the following transformation law holds:

$$\bar{\Gamma}^{k'}_{i'j'} = \sum_{i,j,k} \Gamma^k_{ij} \frac{\partial x^i}{\partial \bar{x}^{i'}} \frac{\partial x^j}{\partial \bar{x}^{j'}} \frac{\partial \bar{x}^{k'}}{\partial x^k} + \sum_k \frac{\partial^2 x^k}{\partial \bar{x}^{i'} \partial \bar{x}^{j'}} \frac{\partial \bar{x}^{k'}}{\partial x^k}.$$

Use this law to check that formula (2.3.2) defines $\frac{\nabla X}{dt}$ independently of choice of local chart.

4 Let M be a Riemannian manifold of dimension n. Given $p \in M$, prove that there exists an open neighborhood U of p, and n smooth vector fields $E_1, \ldots, E_n$ defined on U which are orthonormal at each point of U and such that $(\nabla_{E_i} E_j)_p = 0$ for all i, j.

5 Let M be a Riemannian manifold. Suppose X is a smooth vector field along a smooth curve $\gamma : I \to M$. If $\phi : J \to I$ is a diffeomorphism, define the reparametrizations $\eta = \gamma \circ \phi$ and $Y = X \circ \phi$.

a. Show that Y is a smooth vector field along η.

b. Denote by t, s the parameters along γ, η, resp., where $t = \phi(s)$, and prove that

$$\left(\frac{\nabla}{ds} Y\right)(s) = \left(\frac{\nabla}{dt} X\right)(\phi(s))\, \phi'(s)$$

for $s \in J$, namely, $\frac{\nabla X}{ds} = \frac{\nabla X}{dt}\frac{dt}{ds}$.

c. Deduce that the parallelism of a vector field along a curve does not depend on the parametrization.

6 Let M be a Riemannian manifold. The goal of this exercise is to characterize the curves on M that are geodesics up to a reparametrization.

a. Assume $\gamma : I \to M$ is a geodesic, $\phi : J \to I$ is a diffeomorphism and $\eta : J \to M$ is given by $\eta = \gamma \circ \phi$. Show that there exists a smooth function $f : J \to \mathbf{R}$ such that $\nabla_{\eta'} \eta' = f \eta'$.

b. Conversely, suppose that $\eta : J \to M$ is a regular curve (i.e. η' is nowhere vanishing) satisfying $\nabla_{\eta'} \eta' = f \eta'$ for some smooth function $f : J \to \mathbf{R}$, and show that there exists a diffeomorphism $\phi : J \to I$ such that $\gamma = \eta \circ \phi^{-1}$ is a geodesic.

7 In this exercise, we describe the geodesics of the real hyperbolic space.

a. Describe the geodesics of $M = \mathbf{R}H^n$ in the hyperboloid model using a reflection argument similar to that used in the case of S^n. Namely, show that the geodesic through $p \in M$ with initial unit speed $v \in T_pM$ is given by $\gamma_v(t) = \cosh t\, p + \sinh t\, v$. Show also that the (unique, up to reparametrization) geodesic joining two points p, $q \in M$ is obtained from the intersection of the 2-plane spanned by p, q in $\mathbf{R}^{1,n}$ with the hyperboloid.

b. Use the result of (a) to describe the geodesics of M in Poincaré's ball and upper half-space models (cf. exercises 3 and 4 of chapter 1).

c. Check that in the case in which $n = 2$, the result of (b) coincides with the result of Example 2.4.8.

8 Consider the Poincaré upper half-plane model $\mathbf{R}^2_+ = \{\, (x, y) \in \mathbf{R}^2 \mid y > 0 \,\}$ with the metric $g = \frac{1}{y^2}\left(dx^2 + dy^2\right)$.

a. Prove that any geodesic of $\mathbf{R}^2_+$ is the fixed point set of some isometry. (Hint: Use Example 2.4.8 and Exercise 5 of chapter 1; conjugate R by appropriate isometries of the form τ_a, h_r.) Such isometries deserve to be called *reflections*. Show that the differential of a reflection at a fixed point p is a reflection of $T_p\mathbf{R}^2_+$ on a straight line.

b. Show that the composition of reflections on two geodesics through the point $p = (0, 1)$ yields an isometry that fixes that point and induces a rotation on the tangent space. Show also that any rotation of $T_p\mathbf{R}^2_+$ arises in this way. Deduce that the isometry group of $\mathbf{R}^+_2$ acts transitively on the unit tangent bundle (namely, the set of unit tangent vectors).

A Riemannian manifold with the property that its isometry group acts transitively on its unit tangent bundle is called *isotropic*.

9 Let M be a smooth manifold equipped with a connection ∇. If $\gamma : (-\epsilon, \epsilon) \to M$ is a smooth curve and X is a smooth vector field along γ, prove the following formula:

$$\left(\frac{\nabla}{dt}X\right)_0 = \lim_{t\to 0} \frac{P^{\gamma}_{0,t}X(t) - X(0)}{t}.$$

(Hint: Write X as a linear combination of the vectors in a parallel frame along γ.)

10 Let M be a Riemannian manifold and consider its Levi-Cività connection ∇. If X is a smooth vector field on M and $\{\varphi_t\}$ denotes its local flow, and $v \in T_pM$, prove the following formula:

$$\nabla_v X = \frac{\nabla}{dt}\bigg|_{t=0} d(\varphi_t)_p v.$$

(Hint: Use the first identity in Proposition 2.6.2 in order to commute two different derivatives.)

11 Let X be a Killing field on a Riemannian manifold M. Prove that if p is a critical point of the function $f = ||X||^2$, then the integral curve of X through p is a geodesic.

12 Let G be a Lie group equipped with a bi-invariant metric. Show that the left-invariant vector fields and the right invariant vector fields are Killing fields.

2.10 Additional notes

§1 The development of the idea of connection presented here, usually called an *affine connection*■1■, took some time to evolve to that form. Starting around 1868, Elwin Christoffel became interested in the theory of invariants and wrote six papers on that topic. In these, he introduced the Christoffel symbols and solved the local equivalence problem for quadratic differential forms by essentially introducing the Riemann-Christoffel curvature tensor. These results influenced Gregorio Ricci-Curbastro in Padua to begin his investigations in 1884 on quadratic differential forms. In four papers between 1888 and 1892, Ricci-Curbastro exposed the technique of absolute differential calculus, a new invariant formalism originally constructed to deal with the transformation theory of partial differential equations, which he used to study the transformation theory of quadratic differential forms. A pupil of him, Tulio Levi-Cività, wrote a dissertation, published in 1893, where he developed the calculus of tensors including covariant differentiation, building on ideas from Ricci-Curbastro and Lie's then recently appeared theory of transformation groups. In 1900, Ricci (using this name for the first time instead of his full name) jointly with Levi-Cività published a fundamental paper [RL00] in which preface they state:

■1■?

"The algorithm of absolute differential calculus, the instrument matériel of the methods ... can be found complete in a remark due to Christoffel. But the methods themselves and the advantages they offer have their raison d'être and their source in the intimate relationships that join them to the notion of an n-dimensional variety, which we owe to the brilliant minds of Gauss and Riemann. ... Being thus associated in an essential way with V^n, it is the natural instrument of all those studies that have as their subject such a variety, or in which one encounters as a characteristic element a positive quadratic form of the differentials of n variables or of their derivatives."

When in 1915 Albert Einstein used tensor calculus to explain his theory of relativity, Levi-Cività initiated and kept mathematical correspondence with him until 1917. In that year, inspired by Einstein's general theory of relativity, Levi-Cività made what is probably his most important contribution to mathematics: the introduction of the concept of parallel displacement. His book [Lev05] on absolute diferential calculus, originally a collection of lecture notes in Italian, also contains applications to general relativity.

Soon it was realized that connections existed independently of the Riemannian metric. Between the years of 1918 and 1923, Hermann Weyl's efforts towards the unification of electromagnetism and gravitation brought in new ideas and placed the concept of parallel displacement of a tangent vector at the base of the definition of an affine connection on a smooth manifold. Tensor calculus was systematized by Jan Schouten (who discovered the idea of parallel displacement independently in 1918) in his book *Ricci-Kalkül* in 1924 (entirely rewritten in 1954). At the same time, Élie Cartan introduced in the 1920's projective and conformal connections and, more generally, a new concept of a connection on a manifold. However, at that time, Cartan faced difficulty trying to express notions for which there was no truly suitable language. In [Ehr51], Charles Ehresmann gave a rigorous global definition of a Cartan connection as a special case of a more general notion of connection on a principal bundle, today called an *Ehresmann connection* or simply a *connection*, which is mostly considered to be the definitive one. The axiomatic approach to affine connections that we use in this book is due to Jean-Louis Koszul (cf. [Nom54]). For more details on the history of connections, see the introduction of [Str34]. For the general theory of connections on principal bundles, see [KN96].

§2 The idea of parallel displacement is a simple though deep notion in geometry. Consider a 2-sphere Σ touching a 2-plane π at a point p. Now let Σ roll over π so that the touching point traces a curve γ in Σ, and let q be the endpoint of γ. Suppose v is a vector tangent to π at p. Of course, there is a unique vector v' which is tangent to π at q and parallel to v in the plane. *The parallelism of Levi-Cività says that v', regarded as vector tangent to Σ at q, is the parallel displacement of v, regarded as a vector tangent to Σ at p, along γ.* More generally, one can replace Σ by a 2-surface and let it roll over π to define the parallel displacement of vectors on Σ.

Completeness

3.1 Introduction

Geodesics of Riemannian manifolds were defined in section 2.4 as solutions to a second order ordinary diferential equation that, in a sense, means that they have acceleration zero, or, so to say, that they are the "straightest" curves in the manifold. On the other hand, the geodesics of Euclidean space are the lines, and it is known that line segments are the shortest curves between its endpoints. One of the goals of this chapter is to propose an alternative characterization of geodesics in Riemannian manifolds as the "shortest" curves in the manifold. As we will see soon, in a general Riemannian manifold we cannot expect this property to hold globally, but only locally.

To begin with, we prove the Gauss lemma and use it to introduce a metric space structure in the Riemannian manifold in order to be able to talk about distances and curves that minimize distance. The proposed characterization as the locally minimizing curves then follows easily from some results of section 2.4. Next, a natural question is how far a geodesic can minimize distance. The appropriate category of Riemannian manifolds in which to consider this question is that of complete Riemannian manifolds, namely, Riemannian manifolds whose geodesics can be extended indefinitely. In this context, we prove our first global result which is the fundamental Hopf-Rinow theorem. Finally, the question of how far a geodesic can minimize distance brings us to the notion of cut-locus.

Throughout this chapter, we let (M, g) denote a *connected* Riemannian manifold.

3.2 The metric space structure

As a preparation for the introduction of the metric space structure, we prove the Gauss lemma and use it to show that the radial geodesics emanating from a point and contained in a normal neighborhood are the shortest curves among the piecewise smooth curves with the same endpoints.

So fix a point $p \in M$. By Proposition 2.4.4, there exist $\epsilon > 0$ and an open neighborhood U of p in M such that $\exp_p : B(0_p, \epsilon) \to U$ is a diffeomorphism. Then we have a diffeomorphism

$$f : (0, \epsilon) \times S^{n-1} \to U \setminus \{p\}, \qquad f(r, v) = \exp_p(rv),$$

where S^{n-1} denotes the unit sphere of (T_pM, g_p). Note that $\gamma_v(t) = f(t, v)$ if $|t| < \epsilon$.

3.2.1 Lemma (Gauss, local version) *The radial geodesic γ_v is perpendicular to the hyperspheres $f(\{r\} \times S^{n-1})$ for $0 < r < \epsilon$. It follows that*

$$f^*g = dr^2 + h_{(r,v)}$$

where $h_{(r,v)}$ is the metric induced on S^{n-1} from $f : \{r\} \times S^{n-1} \to M$.

Proof. For a smooth vector field X on S^{n-1}, we denote by $\tilde{X} = f_* X$ the induced vector field on U. Also, we denote by $\frac{\partial}{\partial r}$ the coordinate vector field on $(0, \epsilon)$ and set $\frac{\tilde{\partial}}{\partial r} = f_* \frac{\partial}{\partial r}$. Next, note that $\gamma_v'(r) = \frac{\tilde{\partial}}{\partial r}|_{f(r,v)}$ and that every vector tangent to $S(p,r) := f(\{r\} \times S^{n-1})$ at $f(r,v)$ is of the form $\tilde{X}|_{f(r,v)}$ for some smooth vector field X on S^{n-1}. In view of that, the problem is reduced to proving that $g(\tilde{X}, \frac{\tilde{\partial}}{\partial r}) = 0$ at $f(r,v)$. With this is mind, we start computing

$$
\begin{aligned}
\frac{d}{dr} g\left(\tilde{X}, \frac{\tilde{\partial}}{\partial r}\right) &= g\left(\nabla_{\frac{\tilde{\partial}}{\partial r}} \tilde{X}, \frac{\tilde{\partial}}{\partial r}\right) + g\left(\tilde{X}, \nabla_{\frac{\tilde{\partial}}{\partial r}} \frac{\tilde{\partial}}{\partial r}\right) \\
&= g\left(\nabla_{\tilde{X}} \frac{\tilde{\partial}}{\partial r}, \frac{\tilde{\partial}}{\partial r}\right) \\
&= \frac{1}{2} \tilde{X}\, g\left(\frac{\tilde{\partial}}{\partial r}, \frac{\tilde{\partial}}{\partial r}\right) \\
&= 0,
\end{aligned}
$$

where we have used the following facts: the compatibility of ∇ with g, $\nabla_{\frac{\tilde{\partial}}{\partial r}} \frac{\tilde{\partial}}{\partial r} = 0$ since γ_v is a geodesic, $\nabla_{\frac{\tilde{\partial}}{\partial r}} \tilde{X} - \nabla_{\tilde{X}} \frac{\tilde{\partial}}{\partial r} = [\frac{\tilde{\partial}}{\partial r}, \tilde{X}] = f_*[\frac{\partial}{\partial r}, X] = 0$ and $g\left(\frac{\tilde{\partial}}{\partial r}, \frac{\tilde{\partial}}{\partial r}\right) = 1$. Now we have that $g(\tilde{X}, \frac{\tilde{\partial}}{\partial r}) = 0$ is constant as a function of $r \in (0, \epsilon)$. Hence

$$
g\left(\tilde{X}, \frac{\tilde{\partial}}{\partial r}\right)\Big|_{f(r,v)} = \lim_{r \to 0} g\left(\tilde{X}, \frac{\tilde{\partial}}{\partial r}\right)\Big|_{f(r,v)} = 0
$$

due to the fact that $\tilde{X}|_{f(r,v)} = d(\exp_p)_{rv}(rX_v)$ goes to 0 as $r \to 0$.

Regarding the last assertion in the statement, the above result shows that in the expression of f^*g there are no mixed terms, namely, no terms involving both dr and coordinates on S^{n-1}, and $g\left(\frac{\tilde{\partial}}{\partial r}, \frac{\tilde{\partial}}{\partial r}\right) = 1$ shows that the coefficient of dr^2 is 1. □

3.2.2 Proposition *Let $p \in M$, and let $\epsilon > 0$ be such that $U = \exp_p(B(0_p, \epsilon))$ is a normal neighborhood of p. Then, for any $x \in U$, there exists a unique geodesic γ of length less than ϵ joining p and x. Moreover, γ is the shortest piecewise smooth curve in M joining p to x, and any other piecewise smooth curve joining p to x with the same length as γ must coincide with it, up to reparametrization.*

Proof. We already know that there exists a unique $v \in T_pM$ with $g_p(v,v)^{1/2} < \epsilon$ and $\exp_p v = x$. Taking γ to be $\gamma_v : [0,1] \to M$, it is clear that the length of γ is less than ϵ.

Next, let η be another piecewise curve joining p to x. We need to prove that $L(\gamma) \leq L(\eta)$, where the equality holds if and only if η and γ coincide, up to reparametrization. Without loss of generality, we may assume that η is defined on $[0,1]$ and that $\eta(t) \neq p$ for $t > 0$. There are two cases:

(a) If η is entirely contained in U, then we can write $\eta(t) = f(r(t), v(t))$ for $t > 0$. In this case, due to the Gauss lemma 3.2.1:

$$\begin{aligned} L(\eta) &= \int_0^1 g_{\eta(t)}(\eta'(t),\eta'(t))^{1/2}\,dt \\ &= \int_0^1 \left(r'(t)^2 + h_{(r(t),v(t))}(v'(t),v'(t)) \right)^{1/2} dt \\ &\geq \int_0^1 |r'(t)|\,dt \\ &\geq |r(1) - \lim_{t\to 0} r(t)| \\ &= L(\gamma). \end{aligned}$$

(b) If η is not contained in U, let

$$t_0 = \inf\{\, t \mid \eta(t) \notin U \,\}.$$

Then, using again the Gauss lemma:

$$L(\eta) \geq L(\eta|_{[0,t_0]}) \geq \int_0^{t_0} |r'(t)|\,dt \geq r(t_0) = \epsilon > L(\gamma).$$

In any case, we have $L(\eta) \geq L(\gamma)$. If $L(\eta) = L(\gamma)$, then we are in the first case and $r'(t) > 0$, $v'(t) = 0$ for all t, so η is a radial geodesic, up to reparametrization. □

For points x, $y \in M$, define

$$d(x,y) = \inf\{\, L(\gamma) \mid \gamma \text{ is a piecewise smooth curve joining } x \text{ and } y \,\}.$$

Note that the infimum in general need not be attained. This happens for instance in the case in which $M = \mathbf{R}^2 \setminus \{(0,0)\}$ and we take $x = (-1,0)$, $y = (1,0)$; here $d(x,y) = 2$, but there is no curve of length 2 joining these points.

3.2.3 Proposition *We have that d is a distance on M, and it induces the manifold topology in M.*

Proof. First notice that the distance of any two points is finite. In fact, since a manifold is locally Euclidean, the set of points of M that can be joined to a given point by a piecewise smooth curve is open. This gives a partition of M into open sets. By connectedness, there must be only one such set.

Next, we remark that $d(x,y) = d(y,x)$, since any curve can be reparametrized backwards. Also, the triangular inequality $d(x,y) \leq d(x,z) + d(z,y)$ holds by juxtaposition of curves, and $d(x,x) = 0$ holds by using a constant curve.

In order to have that d is a distance, it only remains to prove that $d(x,y) > 0$ for $x \neq y$. Choose $\epsilon > 0$ such that $U = \exp_x(B(0_x,\epsilon))$ is a normal neighborhood of x; since $\exp_x : B(0_x,\epsilon) \to U$ is diffeomorphism, by decreasing ϵ, if necessary, we may assume that $y \notin U$. If γ is any piecewise smooth curve joining x to y, and $t_0 = \inf\{\, t \mid \gamma(t) \notin U \,\}$, then $L(\gamma) \geq L(\gamma|_{[0,t_0]}) \geq \epsilon$, where the second inequality is a consequence of Proposition 3.2.2. It follows that $d(x,y) > 0$.

Now that we have d is a distance, we remark that the same Proposition 3.2.2 indeed implies that, in the normal neighborhood U of x, namely for $0 < r < \epsilon$, the distance spheres

$$S(x,r) := \{\, z \in M \mid d(z,x) = r \,\}$$

coincide with the geodesic spheres

$$\{\exp_x(v) \mid g_x(v,v)^{1/2} = r\}.$$

In particular, the distance balls

$$B(x,r) := \{ z \in M \mid d(z,x) < r \}$$

coincide with the geodesic balls $\exp_x(B(0_x, r))$. Since the former make up a system of fundamental neighborhoods of x for the topology of (M, d), and the latter make up a system of fundamental neighborhoods of x for the manifold topology of M, and $x \in M$ is arbitrary, it follows that the topology induced by d coincides with the manifold topology of M. □

Combining results of Propositions 2.4.7 and 3.2.2, we now have the following proposition.

3.2.4 Proposition *Let $p \in M$, and let $\epsilon > 0$ be such that U is an ϵ-totally normal neighborhood of p as in Proposition 2.4.7. Then, for any x, $y \in U$, there exists a unique geodesic γ of length less than ϵ joining x and y; moreover, γ depends smoothly on x and y. Finally, the length of γ is equal to the distance between x and y, and γ is the only piecewise smooth curve in M with this property, up to reparametrization.*

Proof. The first part of the statement is just a paraphrase of Proposition 2.4.7. The second one follows from Proposition 3.2.2. □

We say that a piecewise smooth curve $\gamma : [a,b] \to M$ is *minimizing* if $L(\gamma) = d(\gamma(a), \gamma(b))$.

3.2.5 Lemma *Let $\gamma : [a,b] \to M$ be a minimizing curve. Then the restriction $\gamma|_{[c,d]}$ to any subinterval $[c,d] \subset [a,b]$ is also minimizing.*

Proof. Suppose, on the contrary, that γ is not minimizing on $[c,d]$. This means that there is a piecewise smooth curve η from $\gamma(c)$ to $\gamma(d)$ that is shorter than $\gamma|_{[c,d]}$. Consider the piecewise smooth curve $\zeta : [a,b] \to M$ constructed by replacing $\gamma|_{[c,d]}$ by η, namely,

$$\zeta(t) = \begin{cases} \gamma(t) & \text{if } t \in [a,c], \\ \eta(t) & \text{if } t \in [c,d], \\ \gamma(t) & \text{if } t \in [d,b]. \end{cases}$$

Then ζ is a piecewise smooth curve from $\gamma(a)$ to $\gamma(b)$ and it is clear that ζ is shorter than γ, which is a contradiction. Hence, γ is minimizing on $[c,d]$. □

We can now state the promised characterization of geodesics as the locally minimizing curves.

3.2.6 Theorem (Geodesics are the locally minimizing curves) *A piecewise smooth curve $\gamma : [a,b] \to M$ is a geodesic up to reparametrization if and only if every sufficiently small arc of it is a minimizing curve.*

Proof. Just by continuity, every sufficiently small arc of γ is contained in a ϵ-totally normal neighborhood U of some point of M. But the length of a curve in U of realizes the distance between the endpoints of the curve if and only if that curve is a geodesic, up to reparametrization, by Proposition 3.2.4. Since being a geodesic is a local property, the result is proved. □

Since geodesics are smooth, it follows from Lemma 3.2.5 and Theorem 3.2.6 that a minimizing curve must be smooth.

3.3 Geodesic completeness and the Hopf-Rinow theorem

A Riemannian manifold M is called *geodesically complete* if every geodesic of M can be extended to a geodesic defined on all of $\mathbf{R}$. For instance, $\mathbf{R}^n$ satisfies this condition since its geodesics are lines, but $\mathbf{R}^n$ minus one point does not. A more interesting example is the upper half-plane:

$$\{(x,y) \in \mathbf{R}^2 \,|\, y > 0\}.$$

This manifold is not geodesically complete with respect to the Euclidean metric $dx^2 + dy^2$, but it is so with respect to the hyperbolic metric $\frac{1}{y^2}(dx^2 + dy^2)$ (cf. example 2.4.8 of chapter 2). Of course, an equivalent way of rephrasing this definition is to say that M is geodesically complete if and only if $\exp_p$ is defined on all of T_pM, for all $p \in M$.

We will use the following lemma twice in the proof of the Hopf-Rinow theorem.

3.3.1 Lemma *Let (M,g) be a connected Riemannian manifold. Let x, $y \in M$ be distinct points and let S be the geodesic sphere of radius δ and center x in (M,d). Then, for sufficiently small $\delta > 0$, there exists $z \in S$ such that*

$$d(x,z) + d(z,y) = d(x,y).$$

Proof. If $\delta > 0$ is sufficiently small so that the ball $B(0_x, \delta)$ is contained in an open set where $\exp_x$ is a diffeomorphism onto its image, then $S = \exp_x(S(0_x, \delta))$, where $S(0_x, \delta)$ is the sphere of center 0_x and radius δ in (T_xM, g_x). It will also be convenient to assume that $\delta < d(x,y)$. Since S is compact, there exists a point $z \in S$ such that $d(y,S) = d(y,z)$.

If γ is a piecewise smooth curve from x to y parametrized on $[0,1]$, as $d(x,y) > \delta$, we have that γ meets S at a point $\gamma(t)$, and then

$$\begin{aligned} L(\gamma) &= L(\gamma|_{[0,t]}) + L(\gamma|_{[t,1]}) \\ &\geq d(x,\gamma(t)) + d(\gamma(t),y) \\ &\geq d(x,z) + d(z,y). \end{aligned}$$

This implies that $d(x,y) \geq d(x,z) + d(z,y)$. The thesis now follows from the triangle inequality. □

Historically speaking, it is interesting to notice that the celebrated Hopf-Rinow theorem was only proved in 1931 [HR31]. For ease of presentation, we divide its statement into two parts. The proof of (3.3.2) presented below is due to de Rham [dR73] and is different from the original argument in [HR31].

3.3.2 Theorem (Hopf-Rinow) *Let (M,g) be a connected Riemannian manifold.*

a. Let $p \in M$. If $\exp_p$ is defined on all of T_pM, then any point of M can be joined to p by a minimizing geodesic.

b. If M is geodesically complete, then any two points of M can be joined a minimizing geodesic.

The converse of item (b) in the theorem is false, as can be seen simply by taking M to be an open ball (or any convex subset) of $\mathbf{R}^n$ with the induced metric.

Proof of Theorem 3.3.2. Plainly, it is enough to prove assertion (a) as this assertion implies the other one. So we assume that $\exp_p$ is defined on all of T_pM, and we want to produce a minimizing geodesic from p to a given point $q \in M$. Roughly speaking, the idea of the proof is to start from p with a geodesic in the "right direction", and then to prove that this geodesic eventually reaches q.

By Lemma 3.3.1, for sufficiently small $\delta > 0$, there exists p_0 in a normal neighborhood of p such that

$$d(p, p_0) = \delta \quad \text{and} \quad d(p, p_0) + d(p_0, q) = d(p, q).$$

Let $v \in T_pM$ be the unit vector such that $\exp_p(\delta v) = p_0$, and consider $\gamma(t) = \exp_p(tv)$. We have that γ is a geodesic defined on all of $\mathbf{R}$. We will prove that $\gamma(d(p,q)) = q$.

Let $I = \{\, t \in [0, d(p,q)] \mid d(p,q) = t + d(\gamma(t), q)\,\}$. We already know that 0, $\delta \in I$, so I is nonempty. Let $T = \sup I$. Since the distance $d : M \times M \to \mathbf{R}$ is a continuous function, I is a closed set, and thus contains T. Note that the result will follow if we can prove that $T = d(p,q)$. So suppose that $T < d(p,q)$. Then we can apply Lemma 3.3.1 to the points $\gamma(T)$ and q to find $\epsilon > 0$ and $q_0 \in M$ such that

$$d(\gamma(T), q_0) = \epsilon \quad \text{and} \quad d(\gamma(T), q_0) + d(q_0, q) = d(\gamma(T), q). \tag{3.3.3}$$

Hence

$$\begin{aligned} d(p, q_0) &\geq d(p,q) - d(q_0, q) \\ &= d(p,q) - \big(\, d(\gamma(T), q) - d(\gamma(T), q_0)\,\big) \\ &= \big(\, d(p,q) - d(\gamma(T), q)\,\big) + d(\gamma(T), q_0) \\ &= T + \epsilon, \end{aligned} \tag{3.3.4}$$

since $T \in I$. Let η be the unique unit speed minimizing geodesic from $\gamma(T)$ to q_0. Since the concatenation of $\gamma|_{[0,T]}$ and η is a piecewise smooth curve of length $T + \epsilon$ joining p to q_0, it follows from estimate (3.3.4) that $d(p, q_0) = T + \epsilon$. Now the concatenation is a minimizing curve, so by Lemma 3.2.5 and Theorem 3.2.6 it must be a geodesic, thence, smooth. Due to the uniqueness of geodesics with given initial conditions, η must extend $\gamma|_{[0,T]}$ as a geodesic, and therefore $\gamma(T+\epsilon) = \eta(\epsilon) = q_0$. Using this and equations (3.3.3), we finally get that

$$d(q, \gamma(T+\epsilon)) + T + \epsilon = d(q, q_0) + d(\gamma(T), q_0) + T = d(\gamma(T), q) + T = d(p,q),$$

and this implies that $T + \epsilon \in I$, which is a contradiction. Hence the supposition that $T < d(p,q)$ is wrong and the result follows. ☐

3.3.5 Theorem (Hopf-Rinow) *Let (M, g) be a connected Riemannian manifold. Then the following assertions are equivalent:*

a. (M, g) is geodesically complete.
b. For every $p \in M$, $\exp_p$ is defined on all of T_pM.
c. For some $p \in M$, $\exp_p$ is defined on all of T_pM.
d. Every closed and bounded subset of (M, d) is compact.
e. (M, d) is complete as a metric space.

Proof. The assertions that (a) implies (b) and (b) implies (c) are obvious. We start the proof showing that (c) implies (d). Let K be a closed and bounded subset of M. Since K is bounded, there exists $R > 0$ such that $\sup_{x \in K}\{d(p,x)\} < R$. For every $q \in K$, there exists a minimizing geodesic γ from p to q by assumption and the first part of Theorem 3.3.2. Note that $L(\gamma) = d(p,q) < R$. This shows that $K \subset \exp_p(\overline{B(0_p, R)})$. Now K is a closed subset of compact set and thus compact itself.

The proof that (d) implies (e) is a general argument in the theory of complete metric spaces. In fact, any Cauchy sequence in (M, d) is bounded, hence contained in a closed ball, which must be compact by (d). Therefore the sequence admits a convergent subsequence, and thus it is convergent itself, proving (e).

Finally, let us show that (e) implies (a). This is maybe the most relevant part of the proof of this theorem. So assume that γ is a geodesic of (M, g) parametrized with unit speed. The maximal interval of definition of γ is open by Theorem 2.4.2 on the local existence and uniqueness of solutions of second order differential equations; let it be (a, b), where $a \in \mathbf{R} \cup \{-\infty\}$ and $b \in \mathbf{R} \cup \{+\infty\}$.

We claim that γ is defined on all of $\mathbf{R}$. Suppose, on the contrary, that $b < +\infty$. Choose a sequence (t_n) in (a, b) such that $t_n \nearrow b$. Since

$$d(\gamma(t_m), \gamma(t_n)) \leq L(\gamma|_{[t_m,t_n]}) = t_n - t_m$$

for $n > m$, the sequence $(\gamma(t_n))$ is a Cauchy sequence and thus converges to a point $p \in M$ by (e). Let U be a totally normal neighborhood of p given by Proposition 2.4.7 such that every unit speed geodesic starting at a point in U is defined at least on the interval $(-\epsilon, \epsilon)$, for some $\epsilon > 0$. Choose n so that $|t_n - b| < \frac{\epsilon}{2}$ and $\gamma(t_n) \in U$. Then $t_n + \epsilon > b + \frac{\epsilon}{2}$ and the geodesic γ can be extended to $(a, t_n + \epsilon)$, which is a contradiction. Hence $b = +\infty$. Similarly, one shows that $a = -\infty$, and this finishes the proof of the theorem. ☐

We call the attention of the reader to the equivalence of statements (a) and (e) in Theorem 3.3.5. Because of it, hereafter we can say unambiguously that a Riemannian manifold is *complete* if it satisfies either one of assertions (a) or (e). The following are immediate corollaries of the Hopf-Rinow theorem.

3.3.6 Corollary *A compact Riemannian manifold is complete.*

Recall that the *diameter* of a metric space (M, d) is defined to be

$$\operatorname{diam}(M) = \sup\{\, d(x, y) \mid x,\ y \in M \,\}$$

3.3.7 Corollary *A complete Riemannian manifold of bounded diameter is compact.*

As an application of the concept of completeness, we prove the following proposition which will be used in Chapter 6.

3.3.8 Proposition *Let $\pi : (\tilde{M}, \tilde{g}) \to (M, g)$ be a local isometry.*

a. If π is a Riemannian covering map and (M, g) is complete, then $(\tilde{M}, \tilde{g})$ is also complete.

b. If $(\tilde{M}, \tilde{g})$ is complete, then π is a Riemannian covering map and (M, g) is also complete.

Proof. (a) Let $\tilde{\gamma}$ be a geodesic in $\tilde{M}$. Then the curve γ in M defined by $\gamma = \pi \circ \tilde{\gamma}$ is a geodesic of M by Proposition 2.8.3. In view of the completeness of M, γ is defined on all of $\mathbf{R}$. Again by Proposition 2.8.3, $\tilde{\gamma}$ is a lifting of γ, so $\tilde{\gamma}$ can be extended to be defined on all of $\mathbf{R}$, proving that $\tilde{M}$ is geodesically complete.

(b) Let $p \in M$. We need to construct an evenly covered neighborhood p in M. Suppose that $\pi^{-1}(p) = \{\, \tilde{p}_i \in \tilde{M} \mid i \in I \,\}$, where I is some index set. We can choose $r > 0$ such that $\exp_p : B(0_p, r) \to B(p, r)$ is a diffeomorphism, where $B(p, r)$ denotes the open ball in M of center p and radius r. Set $U = B(p, \frac{r}{2})$ and $\tilde{U}_i = B(\tilde{p}_i, \frac{r}{2})$; these are open sets in M, $\tilde{M}$, respectively. Since π is a local isometry by assumption, we have that the diagram

$$\begin{array}{ccc} B(0_{\tilde{p}_i}, \frac{r}{2}) & \xrightarrow{\exp_{\tilde{p}_i}} & \tilde{U}_i \\ {\scriptstyle d\pi_{\tilde{p}_i}}\downarrow & & \downarrow{\scriptstyle \pi} \\ B(0_p, \frac{r}{2}) & \xrightarrow[\exp_p]{} & U \end{array} \tag{3.3.9}$$

is commutative for all i. Next, we use the assumption that $(\tilde{M},\tilde{g})$ is geodesically complete for the first time (it will be used again below). It implies via the Theorem of Hopf-Rinow that any point in $\tilde{U}_i$ can be joined to $\tilde{p}_i$ by a minimizing geodesic, and hence

$$\exp_{\tilde{p}_i}\left(B\left(0_{\tilde{p}_i},\frac{r}{2}\right)\right)=\tilde{U}_i \tag{3.3.10}$$

for all i (note that the direct inclusion is always valid, so we actually used the assumption only to get the reverse inclusion). This, put together with (3.3.9), gives that $\pi(\tilde{U}_i)=U$ for all i, hence

$$\bigcup_{i\in I}\tilde{U}_i\subset\pi^{-1}(U).$$

Since $\exp_p\circ\, d\pi_{\tilde{p}_i}:B(0_{\tilde{p}_i},\frac{r}{2})\to U$ is injective, (3.3.9) and (3.3.10) indeed imply that

$$\pi:\tilde{U}_i\to U$$

is injective; as it is already surjective and a local diffeomorphism, this implies that it is a diffeomorphism. We also claim that the $\tilde{U}_i$ for $i\in I$ are pairwise disjoint. Indeed, if there is a point $q\in\tilde{U}_i\cap\tilde{U}_j$, then

$$d(\tilde{p}_i,\tilde{p}_j)\leq d(\tilde{p}_i,q)+d(q,\tilde{p}_j)<\frac{r}{2}+\frac{r}{2}=r,$$

so $\tilde{p}_j\in B(\tilde{p}_i,r)$. But one sees that π is injective on $B(\tilde{p}_i,r)$ in the same way as we saw that π is injective on $\tilde{U}_i$. It follows that $\tilde{p}_i=\tilde{p}_j$ and hence $i=j$.

It remains to show that $\pi^{-1}(U)\subset\cup_{i\in I}\tilde{U}_i$. Let $\tilde{q}\in\pi^{-1}(U)$. Set $\pi(\tilde{q})=q\in U$. By our choice of r, there is a unique $v\in T_qM$ such that $||v||<\frac{r}{2}$ and $p=\exp_q v$. Let $\tilde{v}=(d\pi_{\tilde{q}})^{-1}(v)\in T_{\tilde{q}}\tilde{M}$. The geodesic $\tilde{\gamma}(t)=\exp_{\tilde{q}}(t\tilde{v})$ is defined on $\mathbf{R}$ since $(\tilde{M},\tilde{g})$ is complete. Now

$$\pi\circ\gamma(1)=\pi\circ\exp_{\tilde{q}}(\tilde{v})=\exp_{\pi(\tilde{q})}((d\pi)_{\tilde{q}}(\tilde{v}))=\exp_q v=p,$$

so $\tilde{\gamma}(1)=\tilde{p}_{i_0}$ for some $i_0\in I$. Since $||\tilde{v}||<\frac{r}{2}$, we have that $\tilde{q}=\tilde{\gamma}(0)\in B(\tilde{p}_{i_0},\frac{r}{2})=\tilde{U}_{i_0}$, as desired.

Now that we know that π is a Riemannian covering, the completeness of M follows from that of $\tilde{M}$ and Proposition 2.8.3. □

We close this section by proving that Killing fields on complete Riemannian manifolds are complete.

3.3.11 Proposition *Let M be a complete Riemannian manifold. Then any Killing field on M is complete as a vector field. It follows that the Lie algebra of Killing fields on M is isomorphic to the Lie algebra of the isometry group of M.*

Proof. Let X be a Killing field on M, and let $\gamma:(a,b)\to M$ be an integral curve of X with $b<+\infty$. In order to prove that X is complete, it suffices to show that γ can be extended to $(a,b]$. In fact formula (2.5.1) implies that $Xg(X,X)=0$, whence $||\gamma'||$ is a constant c. Therefore for t_1, $t_2\in(a,b)$, $t_1<t_2$, we have

$$d(\gamma(t_1),\gamma(t_2))\leq L(\gamma|_{[t_1,t_2]})=c(t_2-t_1).$$

Then it follows from the completeness of M that $\lim_{t\to b-}\gamma(t)$ exists, as desired.

We have proved that Killing fields are infinitesimal generators of (global) one-parameter subgroups of isometries of M. The second assertion follows. □

3.4 Cut locus

Consider the following facts that we have already discussed: every geodesic is locally minimizing (Theorem 3.2.6); a minimizing geodesic remains minimizing when restricted to a subinterval of its domain (Lemma 3.2.5); in a complete Riemannian manifold, the domain of any geodesic can be extended to all of $\mathbf{R}$. In view of this, a natural question can be posed now: how far is a geodesic in a complete Riemannian manifold minimizing? This is the motivation to introduce the concept of cut locus. We start with a lemma.

3.4.1 Lemma *Let M be a connected Riemannian manifold. Let $\gamma : I \to \mathbf{R}$ be a geodesic, where I is an open interval, and let $[a,b] \subset I$.*

a. If there exists another geodesic η of the same length as $\gamma|_{[a,b]}$ from $\gamma(a)$ to $\gamma(b)$, then γ is not minimizing on $[a,b+\epsilon]$ for any $\epsilon > 0$.

b. If (M,g) is complete and no geodesic from $\gamma(a)$ to $\gamma(b)$ is shorter than $\gamma|_{[a,b]}$, then γ is minimizing on $[a,b]$.

Proof. (a) Consider the piecewise smooth curve $\zeta : [a,b+\epsilon] \to M$ defined by

$$\zeta(t) = \begin{cases} \eta(t) & \text{if } t \in [a,b], \\ \gamma(t) & \text{if } t \in [b,b+\epsilon]. \end{cases}$$

Since η and γ are distinct geodesics, ζ is not smooth at $t = b$. It follows that ζ is not minimizing on $[a,b+\epsilon]$. Since γ and ζ have the same length on $[a,b+\epsilon]$, this implies that neither γ is minimizing on this interval.

(b) If M is complete, there exists a minimizing geodesic ζ from $\gamma(a)$ to $\gamma(b)$ by the Hopf-Rinow theorem. Since no geodesic from $\gamma(a)$ to $\gamma(b)$ is shorter than γ, ζ and γ have the same length, so γ is also minimizing. □

Henceforth, in this section, we assume that M is a complete Riemannian manifold. Fix a point $p \in M$. For each unit tangent vector $v \in T_pM$, we define

(3.4.2) $$\rho(v) = \sup\{\, t > 0 \mid d(p, \gamma_v(t)) = t \,\}.$$

Of course, $\rho(v)$ can be infinite. Notice that the set in the right hand side is a closed interval. It is immediate from the definition that γ_v is minimizing on $[0,t]$ if $0 < t \leq \rho(v)$, and γ_v is not minimizing on $[0,t]$ if $t > \rho(v)$. It follows from Lemma 3.4.1 that γ_v is the unique minimizing geodesic from p to $\gamma_v(t)$ if $0 < t < \rho(v)$.

One proves that ρ is a continuous function from the unit tangent bundle UM of M into $(0,+\infty]$ (see exercise 11 in chapter 5); as usual, the topology we are considering in $(0,+\infty]$ is such that a system of local neighborhoods of the point $+\infty$ is given by the complements in $(0,+\infty]$ of the compact subsets of $(0,+\infty)$. By compactness of the unit sphere U_pM of T_pM, it follows that there exists $v_0 \in U_pM$ such that $\rho(v_0) = \inf_{v \in U_pM} \rho(v)$, but it can happen that $\rho(v_0) = +\infty$.

The *injectivity radius at p* is defined to be

$$\mathrm{inj}_p(M) = \{\, \inf \rho(v) \mid v \in T_pM,\ ||v|| = 1 \,\}.$$

It follows that $\mathrm{inj}_p(M) \in (0,+\infty]$. Also, the *injectivity radius* of M is defined to be

$$\mathrm{inj}(M) = \inf_{p \in M} \mathrm{inj}_p(M).$$

One can show that $p \in M \mapsto \mathrm{inj}_p(M) \in (0,+\infty]$ is a continuous function (see exercise 11 of chapter 5).

In the case in which M is compact, its diameter is finite, so no geodesic can be minimizing past $t = \mathrm{diam}(M)$. Hence $\rho(v)$ is finite for every unit vector $v \in T_pM$, and it follows that ρ is bounded and $\mathrm{inj}(M)$ is finite and positive.

The *tangential cut locus* of M at p is defined as the subset of T_pM given by

$$\mathrm{C}_p = \{\, \rho(v)v \in T_pM \mid v \in T_pM,\ ||v|| = 1 \,\}.$$

The *cut locus* of M at p is defined as the subset of M given by

$$\mathrm{Cut}(p) = \exp_p \mathrm{C}_p = \{\, \gamma_v(\rho(v)) \mid v \in T_pM,\ ||v|| = 1 \,\}.$$

We will also consider the star-shaped open subset of T_pM given by

$$\mathrm{D}_p = \{\, tv \in T_pM \mid 0 \leq t < \rho(v),\ v \in T_pM\ , ||v|| = 1 \,\}.$$

Notice that $\partial \mathrm{D}_p = \mathrm{C}_p$ and $\mathrm{inj}_p(M) = d(p, \mathrm{Cut}(p))$ (possibly infinite).

3.4.3 Proposition *Let M be a complete Riemannian manifold. Then, for every $p \in M$, we have a disjoint union*

$$M = \exp_p(\mathrm{D}_p)\dot{\cup}\mathrm{Cut}(p).$$

Proof. Given $x \in M$, by the Hopf-Rinow theorem there exists a minimizing unit speed geodesic γ_v joining p to x, where $v \in T_pM$ and $||v|| = 1$. As γ_v is minimizing on $[0, d(p, x)]$, we have that $\rho(v) \geq d(p, x)$. This implies that $d(p, x)v \in \mathrm{D}_p \cup \mathrm{C}_p$, thence $x = \exp_p(d(p, x)v) \in \exp_p(\mathrm{D}_p) \cup \mathrm{Cut}(p)$ proving that $M = \exp_p(\mathrm{D}_p) \cup \mathrm{Cut}(p)$.

On the other hand, suppose that $x \in \exp_p(\mathrm{D}_p) \cap \mathrm{Cut}(p)$. Then $x \in \exp_p(\mathrm{D}_p)$ means that there exists a minimizing unit speed geodesic $\gamma : [0, a] \to M$ with $\gamma(0) = p$, $\gamma(a) = x$ and γ is minimizing on $[0, a + \epsilon]$ for some $\epsilon > 0$. On the other hand, $x \in \mathrm{Cut}(p)$ means that there exists a minimizing unit speed geodesic $\eta : [0, b] \to M$ with $\eta(0) = p$, $\eta(b) = x$ and η is not minimizing past b. It follows that γ and η are distinct. We reach a contradiction by noting that γ cannot be minimizing past a by Lemma 3.4.1(a). Hence such an x cannot exist, namely, $\exp_p(\mathrm{D}_p) \cap \mathrm{Cut}(p) = \varnothing$. $\square$

We already know that $\exp_p$ is injective on D_p. We will see in Corollary 5.5.4 that $\exp_p$ is a diffeomorphism of D_p onto its image. It follows that, if M is compact, $\exp_p(\mathrm{D}_p)$ is homeomorphic to an open ball in $\mathbf{R}^n$, and M is obtained from $\mathrm{Cut}(p)$ by attaching an n-dimensional cell via the map $\exp_p : \mathrm{C}_p \to \mathrm{Cut}(p)$. In particular, $\mathrm{Cut}(p)$ is a strong deformation retract of $M \setminus \{p\}$:[1]one simply pushes $M \setminus \{p\}$ out to $\mathrm{Cut}(p)$ along the geodesics emanating from p.

3.5 Examples

Empty cut-locus

In the case of $\mathbf{R}^n$ and $\mathbf{R}H^n$, we already know that the geodesics are defined on $\mathbf{R}$, so these Riemannian manifolds are complete (see exercise 7 of chapter 2 for the geodesics of $\mathbf{R}H^n$). We also know that there is a unique geodesic segment joining two given distinct points; since by the Hopf-Rinow theorem there must be a minimizing geodesic joining those two points, that geodesic segment must be the minimizing one. It follows that any geodesic segment is minimizing and hence the cut-locus of any point is empty. This situation will be generalized in chapter 6 (cf. Corollary 6.5.3).

[1]Mention implications for the topology of M.

S^n and $\mathbf{R}P^n$

In the case of S^n, the geodesics are the great circles, so they are defined on $\mathbf{R}$, even if they are all periodic. Therefore S^n is complete. Let $p \in S^n$. A unit speed geodesic γ starting at $\gamma(0) = p$ is minimizing before it reaches the antipodal point $\gamma(\pi) = -p$ because γ is the only geodesic joining p to $\gamma(t)$ for $t \in (0, \pi)$. If $t = \pi + \epsilon$ for some small $\epsilon > 0$, then there is a shorter geodesic η joining p to $\gamma(t)$ which has $\eta'(0) = -\gamma'(0)$. It follows that $\mathrm{Cut}(p) = \{-p\}$.

In the case of $\mathbf{R}P^n$, the geodesics are the projections of the geodesics of S^n under the double covering $\pi : S^n \to \mathbf{R}P^n$. Let $\bar{p} = \pi(p)$. Given two distinct unit speed geodesics γ_1, γ_2 in S^n starting at p, the smallest $t > 0$ for which we can have $\gamma_1(t) = -\gamma_2(t)$ is $t = \pi/2$, namely, the parameter value at which γ_1 and γ_2 reach the equator S^{n-1} of S^n (note that this happens only if $-\gamma_2'(0) = \gamma_1'(0)$). It follows that any unit speed geodesic in $\mathbf{R}P^n$ is minimizing until time $\pi/2$; it is also clear that such a geodesic is not minimizing past time $\pi/2$. It follows that $\mathrm{Cut}(\bar{p})$ is the image of the equator $S^{n-1} \subset S^n$ under π, and is thus isometric to $\mathbf{R}P^{n-1}$.

Rectangular flat 2-tori

The next example we consider is a rectangular 2-torus $\mathbf{R}^2/\Gamma$, where Γ is spanned by an orthogonal basis $\{v_1, v_2\}$ of $\mathbf{R}^2$. We want to describe $\mathrm{Cut}(\bar{p})$, where $\bar{p} = \pi(p)$ for some $p \in \mathbf{R}^2$ and $\pi : \mathbf{R}^2 \to \mathbf{R}^2/\Gamma$ is the projection. For simplicity, assume $p = \frac{1}{2}(v_1 + v_2)$; this entails no loss of generality as $\mathbf{R}^2/\Gamma$ is homogeneous. Then p is the center of the rectangle $\mathcal{R} = \{\, a_1v_1 + a_2v_2 \in \mathbf{R}^2 \mid 0 \leq a_1, a_2 \leq 1\,\}$. If $\bar{x} = \pi(x)$ for some $x \in \mathbf{R}^2$, then the geodesics joining $\bar{p}$ to $\bar{x}$ are exactly the projections of the line segments in $\mathbf{R}^2$ joining p to a point in $x + \Gamma$. It follows that if γ is a line in $\mathbf{R}^2$ starting at p and $\bar{\gamma} = \pi \circ \gamma$ is the corresponding geodesic in $\mathbf{R}^2/\Gamma$ starting at $\bar{p}$, then $\bar{\gamma}$ is minimizing before γ goes out of $\mathcal{R}$, and not afterwards. It follows that $\exp_p(\mathrm{D}_{\bar{p}}) = \pi(\mathrm{int}\,\mathcal{R})$ and $\mathrm{Cut}(\bar{p}) = \pi(\partial\mathcal{R})$ is homeomorphic to the bouquet of two circles $S^1 \vee S^1$.

Riemannian submersions and $\mathbf{C}P^n$

We first describe the behavior of geodesics with regard to Riemannian submersions. Let $\pi : \tilde{M} \to M$ be a Riemannian submersion, and denote by $\mathcal{H}$ the associated horizontal distribution in $\tilde{M}$. A smooth curve in M is called *horizontal* if it is everywhere tangent to $\mathcal{H}$.

3.5.1 Proposition *Let $\pi : \tilde{M} \to M$ be a Riemannian submersion.*

a. We have that π is distance-nonincreasing (or non-expanding*), namely,*

$$d(\pi(\tilde{x}), \pi(\tilde{y})) \leq d(\tilde{x}, \tilde{y})$$

for every $\tilde{x}$, $\tilde{y} \in \tilde{M}$.

b. Let γ be a geodesic of M. Given $\tilde{p} \in \pi^{-1}(\gamma(0))$, there exists a unique locally defined horizontal lift $\tilde{\gamma}$ of γ with $\tilde{\gamma}(0) = \tilde{p}$, and $\tilde{\gamma}$ is a geodesic of $\tilde{M}$.

c. Let $\tilde{\gamma}$ be a geodesic of $\tilde{M}$. If $\tilde{\gamma}'(0)$ is a horizontal vector, then $\tilde{\gamma}'(t)$ is horizontal for every t in the domain of $\tilde{\gamma}$ and the curve $\pi \circ \tilde{\gamma}$ is a geodesic of M of the same length as $\tilde{\gamma}$.

d. If $\tilde{M}$ is complete, then so is M.

Proof. (a) If $\tilde{\gamma}$ is a piecewise smooth curve on $\tilde{M}$ joining $\tilde{x}$ and $\tilde{y}$, then the curve $\pi \circ \tilde{\gamma}$ on M is also piecewise smooth and joins $\pi(\tilde{x})$ and $\pi(\tilde{y})$. Moreover, $L(\pi \circ \tilde{\gamma}) \leq L(\tilde{\gamma})$, because the projection $d\pi : T\tilde{M} \to TM$ kills the vertical components of vectors and preserves the horizontal ones. It follows that $d(\pi(\tilde{x}), \pi(\tilde{y})) \leq d(\tilde{x}, \tilde{y})$.

(b) If γ is constant, there is nothing to be proven, so we can assume that γ is an immersion. Then there is $\epsilon > 0$ such that $N = \gamma(-\epsilon, \epsilon)$ is an embedded submanifold of M. Since π is a submersion, the pre-image $\tilde{N} = \pi^{-1}(N)$ is an embedded submanifold of $\tilde{M}$. Now there is a smooth function $\phi : \tilde{N} \to (-\epsilon, \epsilon)$ such that $\pi(\tilde{x}) = \gamma(\phi(\tilde{x}))$ for every $\tilde{x} \in N$. Using this function, we can define a smooth horizontal vector field on $\tilde{N}$ by setting

$$\tilde{X}_{\tilde{x}} = (d\pi_{\tilde{x}}|_{\mathcal{H}_{\tilde{x}}})^{-1}(\gamma'(\phi(\tilde{x}))). \tag{3.5.2}$$

Given $\tilde{p} \in \pi^{-1}(\gamma(0)) \in \tilde{N}$, let $\tilde{\gamma}$ be the integral curve of $\tilde{X}$ such that $\tilde{\gamma}(0) = \tilde{p}$. Then $\tilde{\gamma}$ is a horizontal curve locally defined around 0, and $\pi \circ \tilde{\gamma} = \gamma$ because of (3.5.2). It remains to see that $\tilde{\gamma}$ is a geodesic. Indeed, using Theorem 3.2.6 and (a) we have that for every t_0 in the domain of $\tilde{\gamma}$, there exists $\delta > 0$ such that

$$L(\tilde{\gamma}|_{[t_0,t_0+h]}) = L(\gamma|_{[t_0,t_0+h]}) = d(\gamma(t_0), \gamma(t_0+h)) \leq d(\tilde{\gamma}(t_0), \tilde{\gamma}(t_0+h))$$

for $0 < h < \delta$, and there is a similar formula for $-\delta < h < 0$. It follows that $\tilde{\gamma}$ is locally minimizing. Since $||\tilde{\gamma}'|| = ||\gamma||$ is a constant, $\tilde{\gamma}$ is already parametrized proportionally to arc-length, hence it is a geodesic.

(c) Let $\tilde{\gamma}$ be a geodesic of $\tilde{M}$ such that $\tilde{\gamma}'(t_0)$ is horizontal for some t_0 in the domain of $\tilde{\gamma}$. Put $\tilde{p} = \tilde{\gamma}(t_0)$ and suppose γ is the geodesic of M with initial conditions $\gamma(t_0) = \pi(\tilde{p})$ and $\gamma'(t_0) = d\pi_{\tilde{p}}(\tilde{\gamma}'(t_0))$. Using (b), we have a horizontal lift $\tilde{\eta}$ of γ with $\tilde{\eta}(t_0) = \tilde{p}$, locally defined around t_0, which is also a geodesic of $\tilde{M}$. Since $\tilde{\gamma}'(t_0)$ and $\tilde{\eta}'(t_0)$ are both horizontal vectors, they coincide and it follows that $\tilde{\gamma}$ and $\tilde{\eta}$ coincide on an open interval around t_0; on this interval, $\tilde{\gamma}'$ is horizontal and $\pi \circ \tilde{\gamma}$ is a geodesic. From the fact that the set of points in the domain of $\tilde{\gamma}$ where $\tilde{\gamma}'$ is horizontal is closed, we deduce that $\tilde{\gamma}'$ is horizontal wherever it is defined and $\pi \circ \tilde{\gamma}$ is a geodesic everywhere. The assertion about the lengths of $\tilde{\gamma}$ and γ plainly follows from the fact that $d\pi_{\tilde{x}} : \mathcal{H}_{\tilde{x}} \to T_{\pi(\tilde{x})}M$ is a linear isometry for $\tilde{x} \in \tilde{M}$.

(d) Let γ be a geodesic of M. By (b), γ admits a horizontal lift $\tilde{\gamma}$ which turns out to be defined on $\mathbf{R}$ due to the completeness of $\tilde{M}$. It follows from (c) that $\pi \circ \tilde{\gamma}$ is a geodesic of M defined on $\mathbf{R}$, which must clearly extend γ. Hence M is complete. □

In the preceding proposition, it can happen that M is complete but $\tilde{M}$ is not. This happens for instance if π is the inclusion of a proper open subset of $\mathbf{R}^n$ into $\mathbf{R}^n$.

Next we turn to the question of describing the cut-locus of $\mathbf{C}P^n$. Consider the Riemannian submersion $\pi : S^{2n+1} \to \mathbf{C}P^n$ where as usual we view S^{2n+1} as the unit sphere in $\mathbf{C}^{n+1}$. Note that $\mathbf{C}P^n$ is complete by Proposition 3.5.1(d). Let $\tilde{p} \in S^{2n+1}$. Since the fibers of π are just the S^1-orbits, the vertical space $\mathcal{V}_{\tilde{p}} = \mathbf{R}(i\tilde{p})$. It follows that the horizontal space $\mathcal{H}_{\tilde{p}} \subset T_{\tilde{p}}S^{2n+1}$ is the orthogonal complement of $\mathbf{R}\{\tilde{p}, i\tilde{p}\} = \mathbf{C}\tilde{p}$ in $\mathbf{C}^{2n+1}$. In view of the proposition, the unit speed geodesics of $\mathbf{C}P^n$ starting at $p = \pi(\tilde{p})$ are of the form $\gamma(t) = \pi(\cos t\tilde{p} + \sin t\tilde{v})$, where $\tilde{v}$ is orthogonal to $\tilde{p}$ and $i\tilde{p}$. It follows that geodesics are defined on $\mathbf{R}$ and periodic of period π.

We agree to retain the above notations and consider another unit speed geodesic starting at $\tilde{p}$, $\eta(t) = \pi(\cos t\tilde{p} + \sin t\tilde{u})$, where $\tilde{u} \in \mathcal{H}_{\tilde{p}}$. Starting at $t = 0$, $\cos t\tilde{p} + \sin t\tilde{v}$ and $\cos t\tilde{p} + \sin t\tilde{u}$ become linearly dependent over $\mathbf{C}$ for the first time at $t = \pi$ (if $\tilde{v}$, $\tilde{u}$ are linearly independent over $\mathbf{C}$) or at $t = \pi/2$ (if $\tilde{v}$, $\tilde{u}$ are linearly dependent over $\mathbf{C}$). This means that γ and η meet for the first time at $t = \pi$ in the first case and at $t = \pi/2$ in the second one. It follows that γ is minimizing on $[0, t_0]$ for $t_0 \leq \pi/2$. By using Lemma 3.4.1, It also follows that γ is not minimizing on $[0, t_0]$ for $t_0 > \pi/2$.

It follows from the discussion in the previous paragraph that $\mathrm{D}_p = B(0_p, \frac{\pi}{2})$ and a typical point in $\mathrm{Cut}(p)$ is of the form $\gamma(\frac{\pi}{2}) = \pi(\tilde{v})$, where $\tilde{v}$ is a unit vector in $\mathcal{H}_{\tilde{p}}$. Since the unit sphere of $\mathcal{H}_{\tilde{p}}$ is isometric to S^{2n-1}, $\mathrm{Cut}(p) = \pi(S^{2n-1})$ turns out to be isometric to $\mathbf{C}P^{n-1}$.

3.6 Additional notes

§1 Let (X, d) be a connected metric space and define the *length* of a continuous curve $\gamma : [a, b] \to X$ to be the supremum of the lengths of all "polygonal paths" inscribed in γ that join $\gamma(a)$ to $\gamma(b)$, namely,

$$L(\gamma) = \sup_P \sum_{i=1}^n d(\gamma(t_{i-1}), \gamma(t_i)),$$

where $P : a = t_0 < t_1 < \cdots < t_n = b$ runs over all subdivisions of the interval $[a, b]$. A curve is called *rectifiable* if its length is finite. Now (X, d) is called a *length space* if the distance between any two points can be "almost" realized by the length of a continuous curve joining the two points, namely, for every x, $y \in X$,

$$d(x, y) = \inf_\gamma L(\gamma),$$

where γ runs over the set of all continuous curves joining x to y. Any piecewise smooth curve in a connected Riemannian manifold is rectifiable and its length in this sense coincides with its length in the sense of (1.3.5). It follows that the underlying metric space of a connected Riemannian manifold is a length space, but length spaces of course form a much larger class of metric spaces involving no a priori differentiability properties. Many concepts and results of Riemannian geometry admit generalizations to the class of length spaces. For instance, geodesics in length spaces are defined to be the continuous, locally minimizing curves, and one proves that if (X, d) is a complete locally compact length space, then any two points are joined by a minimizing geodesic. There is a distance in the set of isometry classes of compact metric spaces called the *Gromov-Hausdorff distance* which turns it into a complete metric space itself (for noncompact spaces, a slightly more general notion of distance is used), and length spaces form a closed subset in this topology. In this sense, length spaces appear as limits of Riemannian manifolds. For an introduction to general length spaces, see [BBI01].

§2 Next, we give an interesting class of examples of length spaces. Namely, one starts with a connected Riemannian manifold (M, g) of dimension n equipped with a smooth distribution $\mathcal{D}$ of dimension k, where $1 < k < n$, and, for x, $y \in M$, declares $d(x, y) = \inf_\gamma L(\gamma)$ where the infimum is taken over the piecewise smooth curves γ joining x to y such that γ' *is tangent to* $\mathcal{D}$ *whenever defined*. If $\mathcal{D}$ is sufficiently generic, in the sense that iterated brackets of arbitrary length of locally defined sections of $\mathcal{D}$ span TM at every point, then one shows that d is finite and (M, d) is a length space. Note that in this definition we have only used the restriction of g to the sections of $\mathcal{D}$. A triple $(M, \mathcal{D}, g)$ where M is a smooth manifold, $\mathcal{D}$ is a bracket-generating smooth distribution as above and g is a smoothly varying choice of inner products on the fibers of $\mathcal{D}$ is called a *sub-Riemannian manifold*, and the associated length space (M, d) is called a *Carnot-Carathéodory space*; such spaces appear for instance in mechanics with non-holonomic constraints and geometric control theory. A very interesting feature of a Carnot-Carathéodory space is that its Hausdorff dimension is always stricly bigger than its manifold dimension. For further reading about sub-Riemannian geometry, we recommend [BR96, Mon02].

3.7 Exercises

1 Let (M, g) be a connected Riemannian manifold and consider the underlying metric space structure (M, d). Prove that any isometry f of (M, g) is *distance-preserving*, that is, it satisfies the condition that $d(f(x), f(y)) = d(x, y)$ for every x, $y \in M$.

2 Describe the isometry group G of $\mathbf{R}^n$:

a. Show that G is generated by orthogonal transformations and translations.
b. Show that G is isomorphic to the semidirect product $\mathbf{O}(n) \ltimes \mathbf{R}^n$, where

$$(B, w) \cdot (A, v) = (BA, Bv + w)$$

for A, $B \in \mathbf{O}(n)$ and v, $w \in \mathbf{R}^n$.
(Hint: Use the result of the previous exercise.)

3 Prove that every isometry of the unit sphere S^n of Euclidean space $\mathbf{R}^{n+1}$ is the restriction of a linear orthogonal transformation of $\mathbf{R}^{n+1}$. Deduce that the isometry group of S^n is isomorphic to $\mathbf{O}(n+1)$. What is the isometry group of $\mathbf{R}P^n$?

4 Prove that every isometry of the hyperboloid model of $\mathbf{R}H^n$ is the restriction of a linear Lorentzian orthochronous transformation of $\mathbf{R}^{1,n}$. Deduce that the isometry group of $\mathbf{R}H^n$ is isomorphic to $\mathbf{O}^+(1, n)$.

5 A *ray* in a complete Riemannian manifold M is a unit speed geodesic $\gamma : [0, +\infty) \to \mathbf{R}$ such that $d(\gamma(0), \gamma(t)) = t$ for all $t \geq 0$. We say that the ray γ *emanates from* $\gamma(0)$.

Let M be a complete Riemannian manifold and assume that M is noncompact. Prove that, for every $p \in M$, there exists a ray γ emanating from p.

6 A *line* in a complete Riemannian manifold M is a unit speed geodesic $\gamma : \mathbf{R} \to M$ such that $d(\gamma(t), \gamma(s)) = |t - s|$ for all t, $s \geq 0$. Also, M is called *connected at infinity* if for every compact set $K \subset M$ there is a compact set $C \supset K$ such that any two points in $M \setminus C$ can be joined by a curve in $M \setminus K$. If M is not connected at infinity, we say that M is *disconnected at infinity.*

Let M be a complete Riemannian manifold and assume that M is noncompact and disconnected at infinity. Prove that M contains a line.

7 Prove that the following assertions for a Riemannian manifold M are equivalent:
a. M is complete.
b. There exists $p \in M$ such that the function $x \mapsto d(p, x)$ is a proper function on M.
c. For every $p \in M$, the function $x \mapsto d(p, x)$ is a proper function on M.

8 A smooth curve $\gamma : I \to M$ in a Riemannian manifold M defined on an interval $I \subset \mathbf{R}$ is said to be *divergent* if the image of γ does not lie in any compact subset of M.

Prove that a Riemannian manifold is complete if and only if every divergent curve in M has infinite length.

9 Prove that on any smooth manifold a complete Riemannian metric can be defined.

10 Let M be a smooth manifold with the property that it is complete with respect to any Riemannian metric in it. Prove that M must be compact. (Hint: Use the results of exercises 5 and 8.)

11 Describe the cut locus of a point in an hexagonal flat 2-torus. Note that its homeomorphism type is different from that of the cut locus of a point in a rectangular flat 2-torus (compare Examples 3.5).

12 Let M_i be complete Riemannian manifolds, where $i = 1, 2$.

a. Show that the product Riemannian manifold $M_1 \times M_2$ is also complete.
b. Let $p_i \in M_i$, where $i = 1, 2$. Show that the cut locus of (p_1, p_2) in $M_1 \times M_2$ is given by $\big(\mathrm{Cut}(p_1) \times M_2\big) \cup \big(M_1 \times \mathrm{Cut}(p_2)\big)$.

13 A Riemannian manifold M is called *homogeneous* if given any two points of M there exists an isometry of M that maps one point to the other.

Prove that a homogeneous Riemannian manifold is complete.

14 A Riemannian manifold M is called *two point-homogeneous* if given any two equidistant pairs of points of M there exists an isometry of M that maps one pair to the other.

Prove that a Riemannian manifold is two point-homogeneous if and only if it is isotropic.

15 Let f, $g : M \to N$ be local isometries between Riemannian manifolds where M is connected. Assume there exists $p \in M$ such that $f(p) = g(p) = q$ and $df_p = dg_p : T_pM \to T_qN$. Prove that $f = g$. (Hint: Show that the set of points of M where f and g coincide up to first order is closed and open.)

16 Let $\gamma : (a, b) \to M$ be a smooth curve in a Riemannian manifold M. Prove that

$$||\gamma'(t)|| = \lim_{h \to 0} \frac{d\big(\gamma(t+h), \gamma(t)\big)}{|h|}$$

for $t \in (a, b)$. (Hint: Use a normal neighborhood of $\gamma(t)$.)

17 Let (M, g) and (M', g') be Riemannian manifolds, and let d and d' be the associated distances, respectively. Show that a distance-preserving map $f : M \to M'$ (cf. exercise 1) is smooth and a local isometry. (Hint: use a normal neighborhood for the smoothness and exercise 16 to prove it is a local isometry.) Conclude that if f is in addition surjective, then it is a global isometry.

18 Let M be a compact Riemannian manifold of dimension at least two. Prove that the following assertions are equivalent:
a. M is simply-connected;
b. $\mathrm{Cut}(p)$ is simply-connected for all $p \in M$;
c. $\mathrm{Cut}(p)$ is simply-connected for some $p \in M$.

19 Let $\pi : M \to N$ be a smooth submersion and fix a complementary subbundle $\mathcal{H}$ to the vertical bundle $\mathcal{V} = \ker d\pi$ in TM. Prove that any smooth curve in N locally admits a horizontal smooth lift. (Hint: Work in a coordinate system on which π has the standard form of submersions, and express the condition that a smooth curve in M is the horizontal lift of a given smooth curve in N as a system of linear ordinary differential equations.)

CHAPTER 4

Curvature

4.1 Introduction

The curvature of a plane curve is the measure of change of the direction of the curve. Assuming the curve parametrized by arc-length and expressing this direction as a unit tangent vector along the curve exhibits the (unsigned) curvature as the modulus of the second derivative of the curve. In the case of a surface in $\mathbf{R}^3$, Gauss had already shown how to measure curvature: this is the rate of change of the normal direction of the surface. Locally, one chooses a unit normal vector field and differentiates it at a point as a map into the unit sphere. Since the surface is two-dimensional, the result is now a map, namely a linear endomorphism of the tangent space at that point. This turns out to be symmetric, hence diagonalizable over $\mathbf{R}$. Its eigenvalues are called the principal curvatures λ_1 and λ_2. They represent the extreme values of the curvatures of the plane curves given by the normal sections to the surface. Equivalently, one can look at $2H = \lambda_1 + \lambda_2$ and $K = \lambda_1\lambda_2$. The second expression is called the Gaussian curvature and, according to Gauss' celebrated *theorema egregium*, has an intrinsic meaning in the sense that it can be expressed solely in terms of the coefficients of the metric in a coordinate system.

Riemann generalized Gauss' results and explained how to define the curvature of a Riemannian manifold M. Here the dimension of M is at least two, so we start by selecting a 2-plane E contained in T_pM. Exponentiating a small neighborhood of 0_p in E gives a piece of surface S through p contained in M. The curvature of M at E is defined to be the Gaussian curvature of S at p. This gives the sectional curvature function.

As it is, this definition cannot be very useful: it is difficult to compute and, especially, it does not reflect relations between the sectional curvatures of neighboring planes. After Riemann, the matter took a few decades more of study to be settled, until tensor calculus entered the scene.

Throughout this chapter, (M, g) denotes a Riemannian manifold and ∇ denotes its Levi-Cività connection.

4.2 The Riemann-Christoffel curvature tensor

The *curvature tensor* is the tri-linear map $R : \Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ given by

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

It is an easy consequence of the Leibniz rule for ∇ that R is $C^\infty(M)$-linear on each argument. As in the case of connections, this suffices to show that the value of $R(X, Y)Z$ at p depends only on X_p, Y_p, and Z_p. Hence we have a tri-linear map

$$R_p : T_pM \times T_pM \times T_pM \to T_pM.$$

The following are the fundamental symmetries of this map.

4.2.1 Proposition (algebraic properties of the curvature tensor) *We have that*

a. $R(X,Y)Z = -R(Y,X)Z$
b. $\langle R(X,Y)Z, W\rangle = -\langle R(X,Y)W, Z\rangle$
c. $\langle R(X,Y)Z, W\rangle = \langle R(Z,W)X, Y\rangle$
d. $R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$ *(first Bianchi identity)*

for every $X, Y, Z, W \in \Gamma(TM)$.

Proof. (a) This is clear from the definition.

(b) We compute

$$\begin{aligned}\langle R(X,Y)Z,Z\rangle &= \langle \nabla_X\nabla_Y Z, Z\rangle - \langle \nabla_Y\nabla_X Z, Z\rangle - \langle \nabla_{[X,Y]}Z, Z\rangle \\ &= X\langle \nabla_Y Z, Z\rangle - \langle \nabla_Y Z, \nabla_X Z\rangle \\ &\quad - \big(Y\langle \nabla_X Z, Z\rangle - \langle \nabla_X Z, \nabla_Y Z\rangle\big) - \frac{1}{2}[X,Y]\langle Z,Z\rangle \\ &= \frac{1}{2}XY\langle Z,Z\rangle - \frac{1}{2}YX\langle Z,Z\rangle - \frac{1}{2}[X,Y]\langle Z,Z\rangle \\ &= 0,\end{aligned}$$

where we have used several times the compatibility of the Levi-Cività connection with the metric. The identity follows.

(d) We compute

$$\begin{aligned}R(X,Y)Z + R(Y,Z)X + R(Z,X)Y &= \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z \\ &\quad + \nabla_Y\nabla_Z X - \nabla_Z\nabla_Y X - \nabla_{[Y,Z]}X \\ &\quad + \nabla_Z\nabla_X Y - \nabla_X\nabla_Z Y - \nabla_{[Z,X]}Y \\ &= \nabla_X(\nabla_Y Z - \nabla_Z Y) - \nabla_{[X,Y]}Z \\ &\quad + \nabla_Y(\nabla_Z X - \nabla_X Z) - \nabla_{[Y,Z]}X \\ &\quad + \nabla_Z(\nabla_X Y - \nabla_Y X) - \nabla_{[Z,X]}Y \\ &= \nabla_X[Y,Z] - \nabla_{[Y,Z]}X \\ &\quad + \nabla_Y[Z,X] - \nabla_{[Z,X]}Y \\ &\quad + \nabla_Z[X,Y] - \nabla_{[X,Y]}Z \\ &= [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] \\ &= 0,\end{aligned}$$

where we have used the fact that the Levi-Cività connection is torsionless several times, and the Jacobi identity in the last line.

(c) We use (a), (b) and (d) to compute

$$\begin{aligned}\langle R(X,Y)Z,W\rangle &= -\langle R(Y,Z)X,W\rangle - \langle R(Z,X)Y,W\rangle \\ &= \langle R(Y,Z)W,X\rangle + \langle R(Z,X)W,Y\rangle \\ &= -\langle R(Z,W)Y,X\rangle - \langle R(W,Y)Z,X\rangle - \langle R(X,W)Z,Y\rangle - \langle R(W,Z)X,Y\rangle \\ &= 2\langle R(Z,W)X,Y\rangle + \langle R(W,Y)X + R(X,W)Y, Z\rangle \\ &= 2\langle R(Z,W)X,Y\rangle - \langle R(Y,X)W,Z\rangle \\ &= 2\langle R(Z,W)X,Y\rangle - \langle R(X,Y)Z,W\rangle,\end{aligned}$$

which gives the result. $\square$

Let $p \in M$ and let $E \subset T_pM$ be a 2-plane. The *sectional curvature* of M at E is defined to be

$$K(E) = K(x,y) = \frac{-\langle R_p(x,y)x,y\rangle}{||x||^2||y||^2 - \langle x,y\rangle^2},$$

where $\{x,y\}$ is a basis of E. One checks that this expression does not depend on the choice of basis of E as follows. It is very easy to see that $K(y,x)$, $K(\lambda x,y)$ $(\lambda \neq 0)$, $K(x+y,y)$ are all equal to $K(x,y)$. But one can get from $\{x,y\}$ to any other basis of E by performing a number of times the simple transformations

$$\left\{\begin{array}{l} x\mapsto y \\ y \mapsto x\end{array}\right., \quad \left\{\begin{array}{l} x\mapsto \lambda x \\ y \mapsto y\end{array}\right., \quad \left\{\begin{array}{l} x\mapsto x+y \\ y \mapsto y\end{array}\right. .$$

4.2.2 Proposition *We have the following identity*

$$\begin{aligned}&\langle R_p(x,y)z,w\rangle \\ &= \frac{1}{6}\frac{\partial^2}{\partial\alpha\partial\beta}\left(\langle R_p(x+\alpha z, y+\beta w)(x+\alpha z), y+\beta w\rangle - \langle R_p(x+\alpha w, y+\beta z)(x+\alpha w), y+\beta z\rangle\right),\end{aligned}$$

where x, y, z, $w \in T_pM$.

Proof. By direct computation. $\square$

It is important to remark that the identity in the preceding proposition is proved using only the algebraic properties of the curvature tensor. Of course, the next corollary is of an algebraic nature as well.

4.2.3 Corollary *The sectional curvature function $E \mapsto K(E)$ and the metric at a point p determine the curvature tensor at p.*

A Riemannian manifold (M,g) of dimension $n \geq 2$ is said to have *constant curvature* κ if for every point $p \in M$ and every 2-plane $E \subset T_pM$, the sectional curvature at E equals κ. A Riemannian manifold (M,g) of dimension $n \geq 2$ is called *flat* if it has constant curvature κ and $\kappa = 0$. This terminology is consistent with the one introduced in section 1.3: since local isometries must preserve the sectional curvature (see end of this section), a Riemannian manifold locally isometric to Euclidean space must have vanishing sectional curvatures; conversely, we will see in chapter 6 that a Riemannian manifold with vanishing sectional curvatures is locally isometric to Euclidean space. A one-dimensional Riemannian manifold is also called flat, although its tangent spaces do not contain 2-planes, since in this case we have $R \equiv 0$ by Proposition 4.2.1(a). A Riemannian manifold is said to have *positive curvature* (resp. *negative curvature*) if the sectional curvature function is positive (resp. negative) everywhere.

If $\dim M = 2$, then a 2-plane E must coincide with T_pM, and then we have a scalar-valued function $K(p) = K(T_pM)$, which can be shown to coincide with the Gaussian curvature of M in the case in which M is a surface in $\mathbf{R}^3$ equipped with the induced metric (cf. Add. notes §2).

Next, suppose that $\dim M \geq 3$. In this case, we say that M has *isotropic curvature at a point* p if $K(E) = \kappa_p$ for every 2-plane $E \subset T_pM$, where κ_p is a real constant. From the definition of sectional curvature, we have that

$$\langle R_p(x,y)x,y\rangle = -\kappa_p\left(||x||^2||y||^2 - \langle x,y\rangle^2\right)$$

for all $p \in M$ and $x, y \in T_pM$. Set

$$\langle R^0_p(x,y)z, w\rangle = -\langle x,z\rangle\langle y,w\rangle + \langle x,w\rangle\langle y,z\rangle,$$

where $p \in M$ and x, y, z, $w \in T_pM$. Then R^0 is a tensor that has the same symmetries as R. Corollary 4.2.3 implies that

$$R_p = \kappa_p R^0_p. \tag{4.2.4}$$

Obviously, a Riemannian manifold with constant curvature has isotropic curvature at all points. It is a result due to Schur that the converse is true in dimensions at least 3.

4.2.5 Lemma (Schur) *Let M be a connected Riemannian manifold. If M has isotropic curvature at all points and* $\dim M \geq 3$*, then it has constant curvature.*

We will prove the above lemma in section 4.4. Note that the curvature tensor of a Riemannian manifold of constant curvature satisfies identity (4.2.4) where κ_p does not depend on p. We also remark that local isometries must preserve the curvature tensor in the following sense, as is easily seen by using arguments from section 2.5. If $f : M \to N$ is a local isometry between two Riemannian manifolds, then

$$R_{f(p)}(df_p(X_p), df(Y_p))df_p(Z_p) = R_p(X_p, Y_p)Z_p \tag{4.2.6}$$

for every $p \in M$ and every X, Y, $Z \in \Gamma(TM)$. Of course, it also follows that $K(df(E)) = K(E)$ for every 2-plane E contained in T_pM and every $p \in M$.

4.2.7 Remark Let $\varphi : N \to M$ be a smooth map, let X, $Y \in \Gamma(TN)$ be vector fields in N and let $U \in \Gamma(\varphi^*TM)$ be a vector field along φ. Recall the induced connection along φ that was introduced in Proposition 2.6.1. Then one can check that the following identity holds:

$$R(\varphi_*X, \varphi_*Y)U = \nabla^\varphi_X\nabla^\varphi_Y U - \nabla^\varphi_Y\nabla^\varphi_X U - \nabla^\varphi_{[X,Y]}U.$$

4.3 The Ricci tensor and scalar curvature

One can say that the Riemann curvature tensor contains so much information about the Riemannian manifold that it makes sense to consider also some simpler tensors derived from it, and these are the Ricci tensor and the scalar curvature.

The *Ricci tensor* Ric at a point $p \in M$ is the bilinear map $\mathrm{Ric}_p : T_pM \times T_pM \to \mathbf{R}$ given by

$$\mathrm{Ric}_p(x,y) = \operatorname{trace}(v \mapsto -R_p(x,v)y),$$

where x, $y \in T_pM$. Note that the Ricci tensor is defined directly in terms of the curvature tensor without involving the metric. It follows immediately from the symmetries of the curvature tensor given by Proposition 4.2.1 that Ric is symmetric, namely,

$$\mathrm{Ric}_p(x,y) = \mathrm{Ric}_p(y,x)$$

for x, $y \in T_pM$ and $p \in M$. So the Ricci tensor is of the same type as the metric tensor g, and it makes sense to compare the two. An *Einstein manifold* is a Riemannian manifold whose Ricci tensor is proportional to the metric. If $\dim M \geq 3$, it follows from Exercise 4 that the constant of proportionality is independent of the point, and hence the condition is that there exists $\lambda \in \mathbf{R}$ such that

$$\mathrm{Ric} = \lambda\, g.$$

Riemannian manifolds satisfying $\mathrm{Ric} = 0$ are called *Ricci-flat*. Of course, a Riemannian manifold of constant sectional curvature is Einstein, and a flat Riemannian manifold is Ricci-flat.

We can also use the metric to view the Ricci tensor at $p \in M$ as a linear map $T_pM \to T_pM$ by setting

$$\langle \mathrm{Ric}(x), y\rangle = \mathrm{Ric}(x, y).$$

for x, $y \in T_pM$. Then it makes sense to take the trace of Ric: the *scalar curvature* is the smooth function $\mathrm{scal} : M \to \mathbf{R}$ given by

$$\mathrm{scal}(p) = \mathrm{trace}\,\mathrm{Ric}_p,$$

where $p \in M$.

Fix a point $p \in M$ and an orthonormal basis $\{e_1, \ldots, e_n\}$ of T_pM. Then

$$\mathrm{Ric}_p(x, y) = -\sum_{j=1}^{n} \langle R(x, e_j)y, e_j\rangle,$$

where x, $y \in T_pM$. In particular, if x is a unit vector, we can assume that $e_1 = x$ and then

$$\mathrm{Ric}_p(x, x) = \sum_{j=2}^{n} K(x, e_j). \tag{4.3.1}$$

The quadratic form (4.3.1) is sometimes called the *Ricci curvature*; of course, its values on the unit sphere of T_pM completely determine the Ricci tensor at p, and (4.3.1) shows that $\mathrm{Ric}_p(x, x)$ is the (unnormalized) average of the sectional curvatures of the 2-planes containing x. We also have that

$$\mathrm{scal}(p) = \sum_{i=1}^{n} \mathrm{Ric}_p(e_i, e_i) = \sum_{i\neq j} K(e_i, e_j) = 2\sum_{i<j} K(e_i, e_j),$$

and this equation shows that the scalar curvature at p is the (unnormalized) average of the sectional curvatures of the 2-planes in T_pM.

4.4 Covariant derivative of tensors ★

At this juncture, we feel like it is time to discuss how to differentiate tensors on a manifold. If M is a Riemannian manifold, there is a canonical way of differentiating smooth vector fields on M, namely, this is given by the Levi-Cività connection ∇. Viewing vector fields as tensor fields of type $(1, 0)$, we can prove that ∇ naturally extends to connections on all tensor bundles $T^{(r,s)}M$. Denote by $c : T^{(r,s)}M \to T^{(r-1,s-1)}M$ an arbitrary contraction.

4.4.1 Proposition *There is a unique family of connections on the tensor bundles $T^{(r,s)}M$ for r, $s \geq 0$, still denoted by ∇, such that the following conditions hold for $X \in \Gamma(TM)$:*

a. $\nabla_X f = Xf$ *for* $f \in C^\infty(M) = \Gamma(T^{(0,0)}M)$*;*

b. $\nabla_X Y$ *for* $Y \in \Gamma(TM)$ *is the covariant derivative associated to the Levi-Cività connection;*

c. ∇_X *commutes with contractions, that is,* $\nabla_X c(T) = c(\nabla_X T)$ *for* $T \in \Gamma(T^{(r,s)}M)$ *with* r, $s > 0$*;*

d. ∇_X *is a derivation, that is,* $\nabla_X(T \otimes T') = \nabla_X T \otimes T' + T \otimes \nabla_X T'$ *for* $T \in \Gamma(T^{(r,s)}M)$ *and* $T' \in \Gamma(T^{(r',s')}M)$.

Proof. One first proves uniqueness, as follows. Let $X \in \Gamma(TM)$ and assume ∇_X is defined and satisfies the conditions in the statement. Using the same argument as in Subsection 2.2, for an open subset U of M we see that if two tensor fields T, $T' \in \Gamma(T^{(r,s)}M)$ coincide on U then $\nabla_X T$ and $\nabla_X T'$ also coincide on U.

It is now enough to show that $\nabla_X(T|_U)$ is uniquely defined. Write T is a coordinate system $(U, x^1, \dots, x^n)$ as

$$T|_U = \sum a^{i_1\cdots i_r}_{j_1\cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots dx^{j_s},$$

where $a^{i_1\cdots i_r}_{j_1\cdots j_s} \in C^\infty(U)$. The Leibniz rule (d) then gives a formula for $\nabla_X(T|_U)$ is terms of the action of ∇_X on functions, vector fields and 1-forms; the first two cases are taken care by (a) and (b), so we need only show that $\nabla_X\omega$ is uniquely defined for a 1-form ω on M. For that purpose, let $Y \in \Gamma(TM)$ and compute, using (a), (b), (c) and (d):

$$\begin{aligned} \nabla_X\omega(Y) &= c(\nabla_X\omega \otimes Y) \\ &= c(\nabla_X(\omega \otimes Y) - \omega \otimes \nabla_X Y) \\ &= \nabla_X c(\omega \otimes Y) - \omega(\nabla_X Y) \\ &= \nabla_X(\omega(Y)) - \omega(\nabla_X Y), \end{aligned}$$

where c denotes the obvious contraction. Since the last line of this equation is $C^\infty(M)$-linear with respect to Y, yields $\nabla_X\omega$ as a 1-form.

For the existence, one first defines for $\omega \in \Gamma(T^{(0,s)}M)$

$$\begin{aligned} &\nabla_X\omega(X_1, \dots, X_s) \\ &= X(\omega(X_1, \dots, X_s)) - \sum_{i=1}^{s} \omega(X_1, \dots, \nabla_X X_i, \dots, X_s). \end{aligned}$$

Next, for $T \in \Gamma(T^{(r,s)}M)$, note that $T(\omega_1, \dots, \omega_r) \in \Gamma(T^{(0,s)}M)$ for $\omega_i, \dots, \omega_r \in \Gamma(T^*M)$, so we can define

$$\begin{aligned} &\nabla_X T(\omega_1, \dots, \omega_r) \\ &= \nabla_X(T(\omega_1, \dots, \omega_r)) - \sum_{i=1}^{r} T(\omega_1, \dots, \nabla_X\omega_i, \dots, \omega_s). \end{aligned}$$

We leave to the reader to check that this definition satisfies (c) and (d). $\square$

As a first application of Proposition 4.4.1, we view g as a tensor field of type $(0,2)$ and note that the condition that the Levi-Cività connection be compatible with the metric (Proposition 2.2.5(b)) can be restated as simply saying that $\nabla g = 0$, since

$$\nabla_X g(Y, Z) = Xg(Y, Z) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z).$$

This is referred to as the *parallelism of the metric.*

As another application of the Proposition 4.4.1, we prove the *second Bianchi identity* in Proposition 4.4.3 below. Since R is $C^\infty(M)$-linear in each variable, we can view it as a tensor field of type $(1,3)$, namely,

$$\begin{aligned} \Gamma(TM) \otimes \Gamma(TM) \otimes \Gamma(TM) \otimes \Gamma(T^*M) &\to \mathbf{R} \\ (X, Y, Z, \omega) &\to \omega(R(X, Y)Z). \end{aligned}$$

Conversely, $\nabla_X R$, as a tensor of type $(1,3)$, can be viewed as a map

$$\Gamma(TM)\otimes\Gamma(TM)\otimes\Gamma(TM)\to\Gamma(TM).$$

It now follows from the definition of ∇_X acting on $\Gamma(T^{(1,3)}M)$ that we have

$$\nabla_X R(Y,Z)W=\nabla_X(R(Y,Z)W)-R(\nabla_X Y,Z)W-R(Y,\nabla_X Z)W-R(Y,Z)\nabla_X W. \tag{4.4.2}$$

4.4.3 Proposition (Second Bianchi identity) *We have that*

$$\nabla_X R(Y,Z)W+\nabla_Y R(Z,X)W+\nabla_Z R(X,Y)W=0 \tag{4.4.4}$$

for every X, Y, Z, $W\in\Gamma(TM)$.

Proof. Dropping the W in (4.4.2) and using the identity $R(X,Y)=[\nabla_X,\nabla_Y]-\nabla_{[X,Y]}$, we get

$$\begin{aligned}\nabla_X R(Y,Z) &= [\nabla_X,R(Y,Z)]-R(\nabla_X Y,Z)-R(Y,\nabla_X Z)\\ &= [\nabla_X,[\nabla_Y,\nabla_Z]]-[\nabla_X,\nabla_{[Y,Z]}]-R(\nabla_X Y,Z)-R(Y,\nabla_X Z)\\ &= [\nabla_X,[\nabla_Y,\nabla_Z]]-\nabla_{[X,[Y,Z]]}-R(X,[Y,Z])-R(\nabla_X Y,Z)-R(Y,\nabla_X Z).\end{aligned}$$

Summing this formula with the other two obtained by cyclic permutation of (X,Y,Z), we see that the first two terms on the right hand side cancel out because of the Jacobi identity, and invoking the relation $\nabla_X Y-\nabla_Y X=[X,Y]$ also makes remaining terms also disappear. The identity is proved. □

Finally, we use the second Bianchi identity to prove Lemma 4.2.5.

Proof of Lemma 4.2.5. We view $\kappa_p=\kappa(p)$ as a function on M. Note that formula (4.2.4) implies that this function is smooth. We use that formula to get

$$\nabla_X R(Y,Z)W=(X\kappa)R^0(Y,Z)W+\kappa\nabla_X R^0(Y,Z)W.$$

Summing over the cyclic permutations of (X,Y,Z), we have

$$(X\kappa)R^0(Y,Z)W+(Y\kappa)R^0(Z,X)W+(Z\kappa)R^0(X,Y)W=0$$

by an application of the second Bianchi identity (4.4.4) to R and R^0. Let X be an arbitrary unit vector field. As $\dim M\geq 3$, we can select Y, Z so that $\{X,Y,Z\}$ is orthonormal. Also, put $W=Y$. Then

$$X\kappa=0.$$

The connectedness of M implies that κ is constant, as desired. □

4.4.5 Remark The *musical isomorphisms* are defined as follows. For each vector field X on the Riemannian manifold (M,g), one can define the differential 1-form ω given by $\omega(Y)=g(X,Y)$. Note that smoothness of g implies that ω is indeed smooth, and non-degeneracy of g at each point implies that this defines an isomorphism between spaces of sections $\flat:\Gamma(TM)\to\Gamma(T^*M)$, the *flat*, so that $\omega=X^\flat$. The inverse isomorphism is naturally called the *sharp*, denoted $\sharp$, so that $X=\omega^\sharp$. The flat and sharp isomorphisms extend to define isomorphisms $\Gamma(T^{(r,s)}M)\to\Gamma(T^{(r',s')}M)$ for $r+s=r'+s'$ and, as is easily seen, the parallelism of the metric implies that these isomorphisms commute with the covariant derivatives on $\Gamma(T^{(r,s)}M)$ and $\Gamma(T^{(r',s')}M)$. As an example, the curvature tensor R can be viewed as a $(0,4)$ tensor, namely, $R(X,Y,Z,W)=W^\flat(R(X,Y)Z)=g(R(X,Y)Z,W)$.

4.5 Examples

Flat manifolds

Euclidean space is flat, since

$$R(X,Y)Z = X(Y(Z)) - Y(X(Z)) - [X,Y](Z) = 0.$$

Since local isometries must preserve the curvature, it follows that the tori $\mathbf{R}^n/\Gamma$ are also flat.

S^n and $\mathbf{R}P^n$

Since S^n is a Riemannian submanifold of $\mathbf{R}^{n+1}$, for its Levi-Cività connection we have that

$$\nabla_X Y = X(Y) - \langle X(Y), \mathbf{p}\rangle \mathbf{p}, \tag{4.5.1}$$

where X, $Y \in \Gamma(TS^n)$ and we have denoted by $\mathbf{p}$ the position vector. It follows that

$$\begin{aligned} \nabla_X\nabla_Y Z &= X(\nabla_Y Z) - \langle X(\nabla_Y Z), \mathbf{p}\rangle\mathbf{p} \\ &= XY(Z) - \langle XY(Z), \mathbf{p}\rangle\mathbf{p} - \langle Y(Z), X\rangle\mathbf{p} - \langle Y(Z), \mathbf{p}\rangle X \\ &\qquad -\langle XY(Z), \mathbf{p}\rangle\mathbf{p} + \langle XY(Z), \mathbf{p}\rangle\mathbf{p} + \langle Y(Z), X\rangle\mathbf{p} \\ &= XY(Z) - \langle XY(Z), \mathbf{p}\rangle\mathbf{p} + \langle Z, Y\rangle X \end{aligned}$$

where we have used that $\langle Y(Z), \mathbf{p}\rangle = -\langle Z, Y\mathbf{p}\rangle = -\langle Z, Y\rangle$ since $\langle Z, \mathbf{p}\rangle = 0$. Therefore,

$$R(X,Y)Z = \langle Y, Z\rangle X - \langle X, Z\rangle Y. \tag{4.5.2}$$

Comparing with (4.2.4) shows we have proved that S^n has constant curvature 1. Since $\mathbf{R}P^n$ is isometrically covered by S^n, it also has constant curvature 1.

$\mathbf{R}H^n$

Consider the hyperboloid model of $\mathbf{R}H^n$ sitting inside the Lorentzian space $\mathbf{R}^{1,n}$. Although the metric in the ambient space is now Lorentzian, the Levi-Cività connection of $\mathbf{R}H^n$ is given by a formula very similar to (4.5.1), namely, the tangential component of the ambient derivative:

$$\nabla_X Y = X(Y) + \langle X(Y), \mathbf{p}\rangle\mathbf{p}.$$

Indeed, one cheks easily that this formula specifies a connection on $\mathbf{R}H^n$ that satisfies the defining conditions for the Levi-Cività connection. A computation very similar to that in the case of S^n thus gives that

$$R(X,Y)Z = -\langle Y, Z\rangle X + \langle X, Z\rangle Y. \tag{4.5.3}$$

Hence $\mathbf{R}H^n$ has constant curvature -1.

Riemannian products

Let $(M,g) = (M_1, g_1) \times (M_2, g_2)$ be a Riemannian product. It follows immediately from the description of the Levi-Cività connection on M for decomposable vector fields (2.8.1) that the curvature tensor of M is given by

$$R_p(x,y)z = R^1_{p_1}(x_1,y_1)z_1 + R^2_{p_2}(x_2,y_2)z_2,$$

where x, y, $z \in T_pM$ for $p = (p_1, p_2) \in M_1 \times M_2$, $x = x_1 + x_2$, $y = y_1 + y_2$, $z = z_1 + z_2$ are the decompositions relative to the splitting $T_pM = T_{p_1}M_1 \oplus T_{p_2}M_2$, and R^i denotes the curvature tensor of M^i.

In particular,

$$g(R_p(x_1, y_2)x_1, y_2) = g_1(R^1_{p_1}(x_1, 0)x_1, 0) + g_2(R^2_{p_2}(0, y_2)0, y_2) = 0.$$

This shows that a *mixed plane* in M, i.e. a plane with nonzero components in both M_1 and M_2, has sectional curvature equal to zero. It also shows that the product of two positively curved Riemannian manifolds has non-negative curvature.

Riemannian submersions and $\mathbf{C}P^n$

Let $\pi : (\tilde{M}, \tilde{g}) \to (M, g)$ be a Riemannian submersion and consider the splitting $T\tilde{M} = \mathcal{H} \oplus \mathcal{V}$ into the horizontal and vertical distributions. A vector field $\tilde{X}$ on $\tilde{M}$ is called:

- *horizontal* if $\tilde{X}_{\tilde{p}} \in \mathcal{H}_{\tilde{p}}$ for all $\tilde{p} \in \tilde{M}$;
- *vertical* if $\tilde{X}_{\tilde{p}} \in \mathcal{V}_{\tilde{p}}$ for all $\tilde{p} \in \tilde{M}$;
- *projectable* if, for fixed $p \in M$, $d\pi(\tilde{X}_{\tilde{p}})$ is independent of $\tilde{p} \in \pi^{-1}(p)$;
- *basic* if it is horizontal and projectable.

Note that if $\tilde{X}$ is a smooth projectable vector field on $\tilde{M}$, then it defines a smooth vector field X on M by setting $X_p = d\pi(X_{\tilde{p}})$ for any $\tilde{p} \in \pi^{-1}(p)$; in this case, $\tilde{X}$ and X are π-related. It also follows from the definitions that a vertical vector field is projectable and, indeed, a vector field on $\tilde{M}$ is vertical if and only if it is π-related to 0.

If X is a smooth vector field on M, it is clear that there exists a unique basic vector field $\tilde{X}$ on $\tilde{M}$ such that $\tilde{X}$ and X are π-related; the vector field $\tilde{X}$ is necessarily smooth and it is called the *horizontal lift* of X.

4.5.4 Lemma *Let $\tilde{X}$, $\tilde{Y}$ be horizontal lifts of X, $Y \in \Gamma(TM)$, resp., and let $U \in \Gamma(T\tilde{M})$ be a vertical vector field. Then the vector fields $[\tilde{X}, \tilde{Y}] - \widetilde{[X, Y]}$ and $[U, \tilde{X}]$ are vertical.*

Proof. Since U is π-related to 0 and $\tilde{X}$ is π-related to X, we have that $[U, \tilde{X}]$ is π-related to $[0, X] = 0$. A similar argument proves the other assertion. □

The next proposition describes the Levi-Cività connection $\tilde{\nabla}$ of $\tilde{M}$ in terms of the Levi-Cività connection ∇ of M. Denote by $(\cdot)^v$ the vertical component of a vector field on $\tilde{M}$.

4.5.5 Proposition *Let $\pi : (\tilde{M}, \tilde{g}) \to (M, g)$ be a Riemannian submersion. If X, $Y \in \Gamma(TM)$ with horizontal lifts $\tilde{X}$, $\tilde{Y} \in \Gamma(T\tilde{M})$, then*

$$\tilde{\nabla}_{\tilde{X}}\tilde{Y} = \widetilde{\nabla_X Y} + \frac{1}{2}[\tilde{X}, \tilde{Y}]^v.$$

Proof. Apply the Koszul formula (2.2.6) to $\tilde{g}(\tilde{\nabla}_{\tilde{X}}\tilde{Y}, \tilde{Z})$, where $\tilde{Z}$ is the horizontal lift of $Z \in \Gamma(TM)$. Since $d\pi$ restricted to each $\mathcal{H}_{\tilde{p}}$ is a linear isometry onto T_pM for $p = \pi(\tilde{p})$,

$$\tilde{X}_{\tilde{p}}\tilde{g}(\tilde{Y}, \tilde{Z}) = X_p g(Y, Z).$$

Also, by the first assertion of Lemma 4.5.4,

$$\tilde{g}_{\tilde{p}}([\tilde{X}, \tilde{Y}], \tilde{Z}) = g_p([X, Y], Z).$$

Hence

$$\tilde{g}_{\tilde{p}}(\tilde{\nabla}_{\tilde{X}}\tilde{Y},\tilde{Z}) = g_p(\nabla_X Y, Z) = \tilde{g}_{\tilde{p}}(\widetilde{\nabla_X Y},\tilde{Z}). \tag{4.5.6}$$

Next, apply the Koszul formula to $\tilde{g}(\tilde{\nabla}_{\tilde{X}}\tilde{Y},U)$, where $U \in \Gamma(T\tilde{M})$ is vertical. Since $\tilde{g}(\tilde{X},\tilde{Y})$ is constant along the fibers of π, $U\tilde{g}(\tilde{X},\tilde{Y}) = 0$. Using the second assertion of Lemma 4.5.4 yields that

$$\tilde{g}(\tilde{\nabla}_{\tilde{X}}\tilde{Y},U) = \frac{1}{2}\tilde{g}([\tilde{X},\tilde{Y}],U). \tag{4.5.7}$$

The desired result is equivalent to (4.5.6) and (4.5.7). □

The next proposition relates the sectional curvatures of M and $\tilde{M}$.

4.5.8 Proposition *Let $\pi : (\tilde{M},\tilde{g}) \to (M,g)$ be a Riemannian submersion. If X, $Y \in \Gamma(TM)$ is an orthonormal pair with horizontal lifts $\tilde{X}$, $\tilde{Y} \in \Gamma(T\tilde{M})$, then*

$$K(X,Y) = \tilde{K}(\tilde{X},\tilde{Y}) + \frac{3}{4}||[\tilde{X},\tilde{Y}]^v||^2.$$

Proof. We start by observing that for a vertical vector field U on $\tilde{M}$,

$$\tilde{g}(\tilde{\nabla}_{\tilde{X}}U,\tilde{Y}) = -\tilde{g}(U,\tilde{\nabla}_{\tilde{X}}\tilde{Y}) = -\frac{1}{2}\tilde{g}(U,[\tilde{X},\tilde{Y}]^v)$$

by Proposition 4.5.5, and

$$\tilde{g}(\tilde{\nabla}_U\tilde{X},\tilde{Y}) = \tilde{g}(\tilde{\nabla}_{\tilde{X}}U,\tilde{Y}) + \tilde{g}([U,\tilde{X}],\tilde{Y}) = \tilde{g}(\tilde{\nabla}_{\tilde{X}}U,\tilde{Y}),$$

by Lemma 4.5.4. Using these identities and (4.5.5) a few times, we have

$$\begin{aligned}\tilde{\nabla}_{\tilde{X}}\tilde{\nabla}_{\tilde{Y}}\tilde{X} &= \tilde{\nabla}_{\tilde{X}}\left(\widetilde{\nabla_Y X}\right) + \frac{1}{2}\tilde{\nabla}_{\tilde{X}}\left([\tilde{Y},\tilde{X}]^v\right)\\ &= \widetilde{\nabla_X\nabla_Y X} + \frac{1}{2}[\tilde{X},\widetilde{\nabla_Y X}]^v - \frac{1}{2}\tilde{\nabla}_{\tilde{X}}\left([\tilde{X},\tilde{Y}]^v\right),\end{aligned}$$

and

$$\begin{aligned}\tilde{g}(\tilde{\nabla}_{\tilde{X}}\tilde{\nabla}_{\tilde{Y}}\tilde{X},\tilde{Y}) &= \tilde{g}(\widetilde{\nabla_X\nabla_Y X},\tilde{Y}) - \frac{1}{2}\tilde{g}(\tilde{\nabla}_{\tilde{X}}[\tilde{X},\tilde{Y}]^v,\tilde{Y})\\ &= g(\nabla_X\nabla_Y X,Y) + \frac{1}{4}||[\tilde{X},\tilde{Y}]^v||^2\end{aligned}$$

Similarly

$$\tilde{g}(\tilde{\nabla}_{\tilde{Y}}\tilde{\nabla}_{\tilde{X}}\tilde{X},\tilde{Y}) = \tilde{g}(\tilde{\nabla}_{\tilde{Y}}\widetilde{\nabla_X X},\tilde{Y}) = g(\nabla_Y\nabla_X X,Y),$$

and

$$\begin{aligned}\tilde{g}(\tilde{\nabla}_{[\tilde{X},\tilde{Y}]}\tilde{X},\tilde{Y}) &= \tilde{g}(\tilde{\nabla}_{\widetilde{[X,Y]}}\tilde{X},\tilde{Y}) + \tilde{g}(\tilde{\nabla}_{[\tilde{X},\tilde{Y}]^v}\tilde{X},\tilde{Y})\\ &= g(\nabla_{[X,Y]}X,Y) - \frac{1}{2}||[\tilde{X},\tilde{Y}]^v||^2.\end{aligned}$$

It follows that

$$\tilde{g}(\tilde{R}(\tilde{X},\tilde{Y})\tilde{X},\tilde{Y}) = g(R(X,Y)X,Y) - \frac{3}{4}||[\tilde{X},\tilde{Y}]^v||^2,$$

and this clearly implies the desired formula. □

We now apply the above results to the question of computing the sectional curvature of $\mathbf{C}P^n$. Consider as usual the Riemannian submersion $\pi : \tilde{M} = S^{2n+1} \to M = \mathbf{C}P^n$. We will first define a complex structure on each tangent space to M.[1] Since the horizontal space $\mathcal{H}_{\tilde{p}} \subset T_{\tilde{p}}S^{2n+1}$, for $\tilde{p} \in S^{2n+1}$, is the orthogonal complement of $\mathbf{R}\{\tilde{p}, i\tilde{p}\} = \mathbf{C}\tilde{p}$ in $\mathbf{C}^{2n+1}$, it follows that $\mathcal{H}_{\tilde{p}}$ is a complex vector subspace of $\mathbf{C}^{n+1}$. We transfer the complex structure of $\mathcal{H}_{\tilde{p}}$ to T_pM, where $p = \pi(\tilde{p})$, by conjugation with the isometry $d\pi_{\tilde{p}}|_{\mathcal{H}_{\tilde{p}}} : \mathcal{H}_{\tilde{p}} \to T_pM$, namely we set

$$J_p v = d\pi_{\tilde{p}} \circ J_0 \circ (d\pi_{\tilde{p}}|_{\mathcal{H}_{\tilde{p}}})^{-1}(v) = d\pi(i\tilde{v}),$$

where $J_0 : \mathbf{R}^{2n+2} \to \mathbf{R}^{2n+2}$ is the standard complex structure on $\mathbf{R}^{2n+2}$ that allows us to identify $\mathbf{R}^{2n+2} \cong \mathbf{C}^{n+1}$, and $\tilde{v}$ is the horizontal lift of v at $\tilde{p}$. Let us check that J_p is well defined in the sense that if we had started with a different point $\tilde{p}' \in \pi^{-1}(p)$, we would have gotten the same result. Indeed $\tilde{p}' = z\tilde{p}$ for some $z \in S^1$. Denote by $\varphi_z : \mathbf{C}^{n+1} \to \mathbf{C}^{n+1}$ the multiplication by z. Then $\pi \circ \varphi_z = \pi$ which, via the chain rule, yields that $d\pi_{\tilde{p}'} \circ \varphi_z = d\pi_{\tilde{p}}$ and hence

$$\begin{aligned} d\pi_{\tilde{p}'} \circ J_0 \circ (d\pi_{\tilde{p}'}|_{\mathcal{H}_{\tilde{p}'}})^{-1} &= d\pi_{\tilde{p}} \circ \varphi_z \circ J_0 \circ \varphi_{z^{-1}} \circ (d\pi_{\tilde{p}}|_{\mathcal{H}_{\tilde{p}}})^{-1} \\ &= d\pi_{\tilde{p}} \circ J_0 \circ (d\pi_{\tilde{p}}|_{\mathcal{H}_{\tilde{p}}})^{-1}, \end{aligned}$$

since φ_z maps $\mathcal{H}_{\tilde{p}}$ onto $\mathcal{H}_{\tilde{p}'}$. Next, it is clear that

$$J_p^2 = -\mathrm{id}_{T_pM},$$

so J_p introduces on T_pM the structure of a complex vector space. It is also easy to see that J_p is a linear isometry because

$$g(J_pv, J_pw) = \tilde{g}(i\tilde{v}, i\tilde{w}) = \tilde{g}(\tilde{v}, \tilde{w}) = g(v, w),$$

where v, $w \in T_pM$ and $\tilde{v}$, $\tilde{w} \in \mathcal{H}_{\tilde{p}}$ are their corresponding lifts, and we have used the fact that multiplication by i is an isometry of $\mathbf{C}^{n+1}$. Now consider J_p for varying $p \in \mathbf{C}P^n$. If X is a smooth vector field on $\mathbf{C}P^n$, then, plainly, $JX = d\pi(i\tilde{X})$, and this implies that also JX is a smooth vector field on $\mathbf{C}P^n$. Hence J is a smooth tensor field of type $(1, 1)$ on $\mathbf{C}P^n$. Next, we introduce the vertical vector field ξ by putting

$$\xi(\tilde{p}) = \frac{d}{d\theta}\Big|_{\theta=0}(e^{i\theta}\tilde{p}) = i\tilde{p} = J_0(\tilde{p}). \tag{4.5.9}$$

Note that ξ is a smooth, unit vector field on S^{2n+1}. Then $\tilde{X}(\xi) = J_0(\tilde{X}) = i\tilde{X}$, so using the expression of the Levi-Cività connection in S^{2n+1} (4.5.1), we have

$$\begin{aligned} \tilde{\nabla}_{\tilde{X}}\xi &= \tilde{X}(\xi) - \langle \tilde{X}(\xi), \mathbf{p}\rangle \mathbf{p} \\ &= i\tilde{X} - \langle i\tilde{X}, \mathbf{p}\rangle \mathbf{p} \\ &= i\tilde{X}, \end{aligned}$$

[1] For a real vector space V, a *complex structure* is an endomorphism $J : V \to V$ such that $J^2 = -\mathrm{id}_V$. A complex structure J on V allows one to view V as a complex vector space with half the real dimension of V, namely, one puts $(a + ib)v = av + bJv$ for all a, $b \in \mathbf{R}$, $v \in V$. A complex structure on V can exist only if the dimension of V is even (since $(\det J)^2 = (-1)^{\dim V}$), in which case there are many such structures, for the general linear group of V acts on the set of complex strutures by conjugation. Finally, if V is an Euclidean space, a complex structure J on V is called *orthogonal* if J is an orthogonal transformation. The standard complex structure of $\mathbf{R}^{2n}$ is given by $J_0(x, y) = (-y, x)$ for all x, $y \in \mathbf{R}^n$, so that the complex vector space $(\mathbf{R}^{2n}, J_0)$ is isomorphic to $\mathbf{C}^n$ via $(x, y) \mapsto x + iy$.

as $i\tilde{X}$ is tangent to the sphere. Therefore

$$\begin{aligned}\tilde{g}(\xi, [\tilde{X}, \tilde{Y}]^v) &= 2\tilde{g}(\xi, \tilde{\nabla}_{\tilde{X}}\tilde{Y}) && \text{(by Proposition 4.5.5)}\\ &= -2\tilde{g}(\tilde{\nabla}_{\tilde{X}}\xi, \tilde{Y})\\ &= -2\tilde{g}(i\tilde{X}, \tilde{Y})\\ &= -2g(JX, Y).\end{aligned}$$

Since ξ is a unit vector field, in view of Proposition 4.5.8, we finally have that

$$K(X, Y) = 1 + 3\langle JX, Y\rangle^2. \tag{4.5.10}$$

In particular, the sectional curvatures of $\mathbf{C}P^n$ lie between 1 and 4. Further, the sectional curvature of a 2-plane E is 4 (resp. 1) if and only if E is complex (resp. totally real).[2] On the other hand, if we change the metric on $\mathbf{C}P^n$ to the quotient metric coming from the Riemannian submersion $\pi : S^{2n+1}(2) \to \mathbf{C}P^n$ where $S^{2n+1}(2)$ denotes the sphere of radius 2, then its sectional curvatures will lie between $\frac{1}{4}$ and 1 (cf. exercise 2).

For a general even-dimensonal smooth manifold M, a smooth tensor field J of type $(1, 1)$ satisfying $J_p^2 = -\mathrm{id}_{T_pM}$ for all $p \in M$ is called an *almost complex structure*. If J is an almost complex structure on M, a Riemannian metric g on M is called a *Hermitian metric* if J_p is a linear isometry of T_pM with respect to g_p for all $p \in M$. If, in addition, J is parallel ($\nabla J \equiv 0$) with respect to the Levi-Cività connection of (M, g), then (M, g, J) is called an *almost Kähler manifold*.

A *complex manifold* is an even dimensional smooth manifold M admitting a *holomorphic atlas*, namely, an atlas whose transition maps are holomorphic maps between open sets of $\mathbf{C}^n$, after identifying $\mathbf{R}^{2n} \cong \mathbf{C}^n$. It is easy to see that a holomorphic atlas allows one to transfer the complex structure of $\mathbf{R}^{2n}$ to the tangent spaces of M so that a complex manifold automatically inherits a canonical almost complex structure. Not all almost complex structures on a smooth manifold are obtained from a holomorphic atlas in this way and the ones that do are called *integrable*. The celebrated Newlander-Nirenberg theorem supplies a criterium for the integrability of almost complex structures, similar to the Frobenius theorem. An almost Kähler manifold with integrable complex structure is called a *Kähler manifold*. An introduction to the theory of complex manifolds is [Wel08].

We come back to the Riemannian submersion $\pi : S^{2n+1} \to \mathbf{C}P^n$ and the almost complex structure J on $\mathbf{C}P^n$. Note first that $\mathbf{C}^n$ is obviously a complex manifold and indeed a Kähler manifold: for vector fields X, $Y : \mathbf{C}^n \to \mathbf{C}^n$ the Levi-Cività connection $\nabla_X^{\mathbf{C}^n} Y = dY(X)$, so the chain rule yields

$$\nabla_X^{\mathbf{C}^n}(J_0Y) = d(J_0 \circ Y)(X) = dJ_0 \circ dY(X) = J_0\nabla_X^{\mathbf{C}^n} Y$$

and hence $\nabla^{\mathbf{C}^n} J_0 = 0$. Now J_0 restricts to an endomorphism of $\mathcal{H}$ and the Levi-Cività connection of S^{2n+1} is obtained from $\nabla^{\mathbf{C}^n}$ by orthogonal projection, so

$$\tilde{\nabla}_{\tilde{X}}(J_0\tilde{Y}) = J_0\tilde{\nabla}_{\tilde{X}}\tilde{Y}$$

from which it follows that

$$\nabla_X(JY) = J\nabla_X Y,$$

for all X, $Y \in \Gamma(T\mathbf{C}P^n)$. This proves that the almost complex structure of $\mathbf{C}P^n$ is parallel. That $\mathbf{C}P^n$ is a Kähler manifold finally follows from the fact that the transition maps (1.3.4) of the smooth atlas constructed in chapter 1 are holomorphic.

[2] A subspace E of an Euclidean vector space V with orthogonal complex structure J is called *totally real* (resp. *complex*) if $J(E) \perp E$ (resp. $J(E) \subset E$).

Lie groups

Let G be a Lie group equipped with a bi-invariant metric. In this example, we will compute the sectional curvatures of G. Denote by $\mathfrak{g}$ the Lie algebra of G. Any 2-plane E contained in T_gG, $g \in G$, is spanned by X_g, Y_g for some X, $Y \in \mathfrak{g}$, so $K(E) = K(X_g, Y_g)$. Further, since left-translations are isometries, we can write $K(X_g, Y_g) = K(X, Y)$ unambiguously. Next, recall the formula (2.8.8) for the covariant derivative. It yields

$$\nabla_X \nabla_Y X = \frac{1}{2}[X, \nabla_Y X] = \frac{1}{4}[X, [Y, X]] = \frac{1}{4}[[X, Y], X],$$
$$\nabla_Y \nabla_X X = 0,$$
$$\nabla_{[X,Y]} X = \frac{1}{2}[[X, Y], X],$$

hence

$$R(X, Y)X = -\frac{1}{4}[[X, Y], X].$$

Assuming that $\{X, Y\}$ is orthonormal and using (2.8.7), we finally get that

$$K(X, Y) = \frac{1}{4}||[X, Y]||^2.$$

We conclude that G has nonnegative curvature. Let $X \in \mathfrak{g}$ be a unit vector and let $\{E_1, \dots, E_n\}$ be an orthonormal basis of $\mathfrak{g}$ with $E_1 = X$. Due to (4.3.1), we also have

$$\mathrm{Ric}(X, X) = \sum_{j=2}^{n} K(X, E_j) = \frac{1}{4}\sum_{j=2}^{n} ||[X, E_j]||^2.$$

It follows that G has positive Ricci curvature in case its center is discrete. We can also rewrite the preceding equation as

$$\mathrm{Ric}(X, X) = -\frac{1}{4}\sum_{j=2}^{n} g([[X, [X, E_j]], E_j) = -\frac{1}{4}\sum_{j=2}^{n} g(\mathrm{ad}_X^2 E_j, E_j) = -\frac{1}{4}\,\mathrm{trace}\,(\mathrm{ad}_X^2).$$

Thus, by bilinearity and polarization,

$$-4\mathrm{Ric}(X, Y) = \mathrm{trace}\,(\mathrm{ad}X \circ \mathrm{ad}Y) \tag{4.5.11}$$

for every X, $Y \in \mathfrak{g}$.

For a general Lie group G, the right-hand side of equation (4.5.11) defines a bilinear symmetric form $B_{\mathfrak{g}}$ on $\mathfrak{g}$ called the *Killing form* (or *Cartan-Killing form*) of $\mathfrak{g}$, and one easily checks that

$$B_{\mathfrak{g}}(\mathrm{ad}_Z X, Y) + B_{\mathfrak{g}}(X, \mathrm{ad}_Z Y) = 0$$

for every X, Y, $Z \in \mathfrak{g}$. If, in addition, G is compact and the center of $\mathfrak{g}$ is trivial, then one shows that $-B_{\mathfrak{g}}$ is also positive definite [Hel78, Prop. 6.6]. Assuming further that G is connected, it follows by Proposition 2.8.5 and the discussion in chapter 1 that $-B_{\mathfrak{g}}$ induces a bi-invariant metric on G. Hence, in the special case in which the bi-invariant metric on G comes from the Killing form, equation (4.5.11) shows that the Ricci tensor is a multiple of the metric tensor, and G is thus an Einstein manifold.

4.6 Additional notes

§1 We make a small digression into the classical theory of surfaces in $\mathbf{R}^3$, see e.g. [Car76], and prove the following proposition.

4.6.1 Proposition *Let M be a regular surface in $\mathbf{R}^3$ equipped with the induced metric. Then the sectional curvature and the Gaussian curvature of M coincide at each point $p \in M$.*

Proof. Let $\mathbf{x} : U \to M$ be a parametrization, where U is an open subset of $\mathbf{R}^2$. We have that $\{\mathbf{x}_u, \mathbf{x}_v\}$ span the tangent plane to M at each point. The smooth functions $E = \langle \mathbf{x}_u, \mathbf{x}_u \rangle$, $F = \langle \mathbf{x}_u, \mathbf{x}_v \rangle$, $G = \langle \mathbf{x}_v, \mathbf{x}_v \rangle$ are the coefficients of the first fundamental form of M (the induced Riemannian metric). The unit normal vector field is given by

$$N = \frac{\mathbf{x}_u \times \mathbf{x}_v}{||\mathbf{x}_u \times \mathbf{x}_v||}.$$

This defines the Gauss map $N : M \to S^2$. Its differential at $p \in M$ is a symmetric linear map $dN_p : T_pM \to T_pM$ which is represented in the basis $\{\mathbf{x}_u, \mathbf{x}_v\}$ by the matrix

$$\begin{pmatrix} e & f \\ f & g \end{pmatrix}.$$

Using the Christoffel symbols, we can write

$$\begin{aligned} \mathbf{x}_{uu} &= \Gamma^1_{11}\mathbf{x}_u + \Gamma^2_{11}\mathbf{x}_v + eN \\ \mathbf{x}_{uv} &= \Gamma^1_{12}\mathbf{x}_u + \Gamma^2_{12}\mathbf{x}_v + fN \\ \mathbf{x}_{vv} &= \Gamma^1_{22}\mathbf{x}_u + \Gamma^2_{22}\mathbf{x}_v + gN \end{aligned}$$

The sectional curvature of M is given by

$$\begin{aligned} K(\mathbf{x}_u, \mathbf{x}_v) &= \frac{-\langle R(\mathbf{x}_u, \mathbf{x}_v)\mathbf{x}_u, \mathbf{x}_v \rangle}{||\mathbf{x}_u||^2||\mathbf{x}_v||^2 - \langle \mathbf{x}_u, \mathbf{x}_v \rangle^2} \\ &= -\frac{\langle \nabla_{\mathbf{x}_u}\nabla_{\mathbf{x}_v}\mathbf{x}_u - \nabla_{\mathbf{x}_v}\nabla_{\mathbf{x}_u}\mathbf{x}_u, \mathbf{x}_v \rangle}{EG - F^2}, \end{aligned}$$

since $[\mathbf{x}_u, \mathbf{x}_v] = 0$. The Levi-Cività connection ∇ is just the tangential component of the derivative in $\mathbf{R}^3$, so $\nabla_{\mathbf{x}_v}\mathbf{x}_u = (\mathbf{x}_{vu})^\top = \Gamma^1_{12}\mathbf{x}_u + \Gamma^2_{12}\mathbf{x}_v$ and

$$\begin{aligned} \nabla_{\mathbf{x}_u}\nabla_{\mathbf{x}_v}\mathbf{x}_u &= \big((\Gamma^1_{12})_u\mathbf{x}_u + \Gamma^1_{12}\mathbf{x}_{uu} + (\Gamma^2_{12})_u\mathbf{x}_v + \Gamma^2_{12}\mathbf{x}_{uv} \big)^\top \\ &= \big((\Gamma^1_{12})_u + \Gamma^1_{12}\Gamma^1_{11} + \Gamma^2_{12}\Gamma^1_{12} \big)\mathbf{x}_u + \big((\Gamma^2_{12})_u + \Gamma^1_{12}\Gamma^2_{11} + (\Gamma^2_{12})^2 \big)\mathbf{x}_v. \end{aligned}$$

Similarly, one computes that

$$\nabla_{\mathbf{x}_v}\nabla_{\mathbf{x}_u}\mathbf{x}_u = \big((\Gamma^1_{11})_v + \Gamma^1_{11}\Gamma^1_{12} + \Gamma^2_{11}\Gamma^1_{22} \big)\mathbf{x}_u + \big((\Gamma^2_{11})_v + \Gamma^1_{11}\Gamma^2_{12} + \Gamma^2_{11}\Gamma^2_{22} \big)\mathbf{x}_v.$$

It follows from formulas (5) and (5a) in [Car76, section 4.3] that $K(\mathbf{x}_u, \mathbf{x}_v)$ equals the Gaussian curvature of M. We realize that this proof is really a restatement of the proof of the *Theorema Egregium.* In chapter 7, we will present an alternative way of proving this proposition. □

§2 Curvature, in any of its manifestations, is the single most important invariant in Riemannian geometry. It is a local invariant that severely restricts the possibilities for local isometries of a

Riemannian manifold; this is partially reflected in the fact that the group of global isometries of a Riemannian manifold is a finite-dimensional Lie group. At the same time, it is really the presence of curvature that gives rise to the huge variety of non-equivalent Riemannian metrics on a given smooth manifold that we can see. The curvature tensor and its covariant derivatives are indeed the only Riemannian invariants if one demands that they be algebraic invariants stemming from the connection. However, if one requires only tensors that are invariant under isometries — the so-called *natural tensors* — then there is not even hope of achieving a classification without imposing further restrictions [Eps75].

§3 Does the curvature determine the metric? This is a very natural question, and an interesting result of Kulkarni [Kul70] asserts that diffeomorphisms preserving the sectional curvature are isometries if the sectional curvature is not constant and the dimension is bigger than 3. On the other hand, it is important to realize that the curvature tensor, in general, does *not* determine the metric, even given that for $n > 3$ the dimension of the space of (pointwise) curvature tensors $\frac{n^2(n^2-1)}{12}$ is much larger than the dimension of the (pointwise) metric tensors $\frac{n(n-1)}{2}$. Indeed, there are many examples of nonisometric Riemannian manifolds admitting diffeomorphisms that preserve the respective curvature tensors. Of course, the difference between the curvature tensor and the sectional curvature is that the latter involves the metric.

4.7 Exercises

1 Let M be an n-dimensional Riemannian manifold of constant curvature κ. Compute that

$$\mathrm{Ric} = (n-1)\kappa g \quad \text{and} \quad \mathrm{scal} = n(n-1)\kappa.$$

2 Let g and $\bar{g}$ be two Riemannian metrics in the smooth manifold M such that $\bar{g} = \lambda g$ for a constant $\lambda > 0$. Show that the curvature tensor, the sectional curvature, the Ricci tensor and the scalar curvature of the Riemannian manifolds $(M, \bar{g})$ and (M, g) are related by the following equations:

$$\overline{R} = R, \quad \overline{K} = \lambda^{-1} K, \quad \overline{\mathrm{Ric}} = \mathrm{Ric} \quad \text{and} \quad \overline{\mathrm{scal}} = \lambda^{-1}\mathrm{scal}.$$

3 Use the symmetries of the curvature tensor to show that the Ricci tensor determines the curvature tensor in a Riemannian manifold of dimension 3.

4 Let M be a connected Einstein manifold of dimension at least 3. Prove that the constant of proportionality is independent of the point. Deduce Lemma 4.2.5 from this result.

5 Let M be a Riemannian manifold with the property that for any two points p, $q \in M$, the parallel transport map from p to q along a piecewise smooth curve γ joining p to q does not depend on γ. Prove that M must be flat.

6 As a partial converse to the previous exercise, suppose M is a flat manifold, p, $q \in M$, and γ_0, γ_1 are two smooth curves joining p to q. Prove that if γ_0 and γ_1 are smoothly homotopic with the endpoints fixed, then the parallel transport maps from p to q along γ_0 and along γ_1 coincide.

7 Prove that the curvature tensor of $\mathbf{C}P^n$ is

$$R(X,Y)Z = -\langle X, Z\rangle Y + \langle Y, Z\rangle X + \langle X, JZ\rangle JY - \langle Y, JZ\rangle JX + 2\langle X, JY\rangle JZ$$

for vector fields X, Y, Z on $\mathbf{C}P^n$. (Hint: Use formula (4.5.10).)

8 Prove that the curvature tensor and the Ricci tensor of a Kähler manifold (M, g, J) satisfy the following identities:

$$R(X,Y)J = JR(X,Y), \quad R(JX,JY) = R(X,Y) \quad \text{and} \quad \mathrm{Ric}(JX,JY) = \mathrm{Ric}(X,Y),$$

for all vector fields X and Y on M.

9 Prove that the curvature tensor of a Riemannian manifold satisfies the following identities:
a. For tangent vectors x, y, z and w, we have

$$\begin{aligned} 6\langle R(x,y)z,w\rangle &= \langle R(x,y+z)(y+z),w\rangle - \langle R(x,y-z)(y-z),w\rangle \\ &\quad +\langle R(y,x-z)(x-z),w\rangle - \langle R(y,x+z)(x+z),w\rangle \end{aligned}$$

b. For tangent vectors a, b, c, we have

$$4\langle R(a,b)a,c\rangle = \langle R(a,b+c)a,b+c\rangle - \langle R(a,b-c)a,b-c\rangle$$

Deduce an alternative proof of Corollary 4.2.3.

10 Extend the notion of parallel transport along a curve to tensors of type (r, s).

11 Let $\varphi : N \to M$ be a smooth map, let X, $Y \in \Gamma(TN)$ be vector fields in N and let U, $V \in \Gamma(\varphi^*TM)$ be vector fiels along φ. Prove that

$$R(\varphi_*X, \varphi_*Y)U = \nabla^\varphi_X\nabla^\varphi_Y U - \nabla^\varphi_X\nabla^\varphi_Y U - \nabla^\varphi_{[X,Y]}U$$

where R denotes the curvature tensor of M and ∇^φ denotes the induced connection along φ. (Hint: Imitate the argument in the proof of Proposition 2.6.2.)

12 (Riemannian volume) Let (M, g) be an oriented Riemannian manifold of dimension n. Let $\mathcal{E} = (E_1, \ldots, E_n)$ be a positively oriented orthonormal frame on an open subset U (that is, $E_1, \ldots, E_n$ are smooth vector fields defined on U which are orthonormal at each point), and let $(\theta^1, \ldots, \theta^n)$ be the dual coframe of 1-forms on U. Define the n-form $\omega_{\mathcal{E}} = \theta^1 \wedge \cdots \wedge \theta^n$ on U.
a. Prove that for another positively oriented orthonormal frame $\mathcal{E}'$ defined on U' we have $\omega_{\mathcal{E}} = \omega_{\mathcal{E}'}$ on $U \cap U'$. Deduce that there exists a smooth differential form vol_M of degree n on M such that

$$(\mathrm{vol}_M)_p(e_1, \ldots, e_n) = 1$$

for every positively oriented orthonormal basis $e_1, \ldots, e_n$ of T_pM and all $p \in M$. The n-form vol_M is called the *volume form* of (M, g) and the associated measure is called the *Riemannian measure* on M associated to g.
b. Show that for a positively oriented basis $v_1, \ldots, v_n$ of T_pM, we have

$$(\mathrm{vol}_M)_p(v_1, \ldots, v_n) = \sqrt{\det\left(g_p(v_i, v_j)\right)}.$$

Deduce that, in local coordinates $(U, \varphi = (x^1, \ldots, x^n))$,

$$\mathrm{vol}_M = \sqrt{\det(g_{ij})}\, dx^1 \wedge \cdots \wedge dx^n.$$

13 Let (M, g) be an n-dimensional Riemannian manifold.

a. For any smooth function $f : M \to \mathbf{R}$, the *gradient* of f is the smooth vector field $\mathrm{grad} f$ defined by $g((\mathrm{grad} f)_p, v) = df_p(v)$ for all $v \in T_pM$ and all $p \in M$. Prove that

$$\mathrm{grad}(f_1 + f_2) = \mathrm{grad} f_1 + \mathrm{grad} f_2 \quad \text{and} \quad \mathrm{grad}(f_1 f_2) = f_1\, \mathrm{grad} f_2 + f_2\, \mathrm{grad} f_1$$

for all smooth functions f_1, f_2 on M.

b. For any smooth vector field X on M, the *divergence* of X is the smooth function $\mathrm{div}\, X = \mathrm{trace}\,(v \mapsto \nabla_v X)$. Prove that

$$\mathrm{div}\,(X + Y) = \mathrm{div}\, X + \mathrm{div}\, Y \quad \text{and} \quad \mathrm{div}\,(fX) = \langle \mathrm{grad} f, X\rangle + f\, \mathrm{div}\, X$$

for all smooth fuctions f and smooth vector fields X, Y on M.

c. For any smooth function f on M, the *Laplacian* of f is the smooth function $\Delta f = \mathrm{div}\, \mathrm{grad} f$. The function f is called *harmonic* is $\Delta f = 0$. Prove that

$$\Delta(f_1 f_2) = f_1 \Delta f_2 + 2\langle \mathrm{grad} f_1, \mathrm{grad} f_2\rangle + f_2 \Delta f_1$$

for all smooth functions f_1, f_2 on M.

d. For any smooth function f on M, the *Hessian* of f is the $(0, 2)$-tensor $\mathrm{Hess}(f) = \nabla df$. Prove that

$$\mathrm{Hess}(f)(X, Y) = X(Y f) - (\nabla_X Y) f$$

and

$$\mathrm{Hess}(f)(X, Y) = \mathrm{Hess}(f)(Y, X)$$

for all smooth vector fields X, Y on M. Show also that the trace of the Hessian coincides with the Laplacian.

14 (Divergence theorem) Let M be an oriented Riemannian manifold.

a. Prove that for any smooth vector field

$$L_X(dV) = (\mathrm{div} X)\, dV$$

where dV denotes the volume form vol_M. A vector field is called *incompressible* if it is divergence-free. Deduce that a vector field is incompressible if and only if its local flows are volume preserving.

b. Suppose now Ω is a domain in M with smooth boundary and let $\partial\Omega$ be oriented by the outward unit normal ν. Denote the Riemannian volume form of $\partial\Omega$ by dS. Use Stokes' theorem to show that for any compactly supported smooth vector field X on M we have

$$\int_\Omega \mathrm{div}\, X\, dV = \int_{\partial\Omega} \langle X, \nu\rangle\, dS$$

15 (Green identities) Let M be an oriented Riemannian manifold and let Ω be a domain in M as in exercise 14.

a. Prove the "integration by parts formula"

$$\int_\Omega f_1 \Delta f_2\, dV + \int_\Omega \langle \mathrm{grad}\, f_1, \mathrm{grad}\, f_2\rangle\, dV = \int_{\partial\Omega} f_1 \frac{\partial f_2}{\partial \nu}\, dS$$

for any compactly supported smooth functions f_1, f_2 on M. Deduce the *weak maximum principle*: if f is compactly supported and sub- or super-harmonic (i.e. $\Delta f \geq 0$ or $\Delta f \leq 0$) then f is constant. (Hint: first show $\Delta f = 0$; then apply integration by parts to $f = f_1 = f_2$ and $\Omega = M$.)

b. Prove that

$$\int_{\Omega} (f_1 \Delta f_2 - f_2 \Delta f_1) \, dV = \int_{\partial\Omega} \left(f_1 \frac{\partial f_2}{\partial \nu} - f_2 \frac{\partial f_1}{\partial \nu} \right) dS$$

for any compactly supported smooth functions f_1, f_2 on M. Deduce that if f_1 and f_2 are two eigenfunctions of the Laplacian on a compact oriented Riemannian manifold M associated to different eigenvalues λ_1, λ_2, resp., then f_1 and f_2 are orthogonal in the sense that $\int_M f_1 f_2 \, dV = 0$.

Variational calculus

5.1 Introduction

We continue to study the problem of minimization of geodesics in Riemannian manifolds that was started in chapter 3. We already know that geodesics are the locally minimizing curves. Also, long segments of geodesics need not be minimizing, and the study of this phenomenon in complete Riemannian manifolds motivates the definition of cut locus.

Herein we take a different standpoint in that we consider finite segments of curves. Namely, consider a complete Riemannian manifold M. Given two points p, $q \in M$, the Hopf-Rinow theorem ensures the existence of at least one minimizing geodesic γ joining p and q. It follows that γ is a global minimum for the length functional L defined in the space of piecewise smooth curves joining p and q. Of course, the calculus approach to finding global minima of a function is to differentiate it, compute critical points and decide which of them are local minima by using the second derivative. In our case, the apparatus of classical calculus of variations can be applied to carry out this program.

To begin with, we show that the critical points of the length functional in the space of piecewise smooth curves joining p and q are exactly the geodesic segments, up to reparametrization. The main result of this chapter is the Jacobi-Darboux theorem that gives a necessary and sufficient condition for a geodesic segment between p and q to be a local minimum for L. In order to prove this theorem, we introduce Jacobi fields and conjugate points. Finally, we study the relation between the concepts of cut locus and conjugate locus. These results will be generalized in chapter 7, where we will prove the Morse index theorem.

Throughout this chapter, (M, g) denotes a Riemannian manifold.

5.2 The energy functional

Instead of working with the length functional L, we will be working with the energy functional E, which will be defined in a moment. The reason for that is that the critical point theory of E is very much related to the one of L and, from a variational calculus point of view, E is easier to work with than L.

The *energy* of a piecewise smooth curve $\gamma : [a, b] \to M$ is defined to be

$$E(\gamma) = \frac{1}{2}\int_a^b ||\gamma'(t)||^2 \, dt.$$

The factor 1/2 in this expression is a normalization constant and it is not very important.

It is interesting to note that, in contrast to L, E is not invariant under reparametrizations of the curve. On the one hand, this points out the fact that E is not a geometrical invariant like L. On the other hand, this can be seen as an advantage since, as we will soon see, critical points of E come already equipped with a very specific parametrization.

5.2.1 Lemma *Let $\gamma : [a, b] \to M$ be a piecewise smooth curve, and let $\gamma(a) = p$ and $\gamma(b) = q$.*

a. *If γ is minimizing, that is $L(\gamma) = d(p, q)$, then γ is a geodesic, up to reparametrization.*
b. *If γ minimizes the energy in the space of piecewise smooth curves defined on $[a, b]$ and joining p and q, then γ is a minimizing geodesic.*

Proof. (a) If γ is minimizing, then it is locally minimizing (Lemma 3.2.5) and hence a geodesic (Theorem 3.2.6).

(b) In the space of continuous functions $[a, b] \to \mathbf{R}$, consider the scalar product $\langle f, g\rangle = \int_a^b f(t)g(t)\,dt$. The Cauchy-Schwarz inequality says that $\langle f, g\rangle^2 \leq ||f||^2||g||^2$ with the equality holding if and only if $\{f, g\}$ is linearly dependent. Applying this to $f = ||\gamma'||$ and $g = 1$ yields that

$$\left(\int_a^b ||\gamma'(t)||\,dt\right)^2 \leq (b-a)\int_a^b ||\gamma'(t)||^2\,dt,$$

and hence

$$L(\gamma)^2 \leq 2E(\gamma)(b-a) \tag{5.2.2}$$

with the equality holding if and only if γ is parametrized with constant speed. Let η be any piecewise smooth curve defined on $[a, b]$ and joining p and q, and assume that it is parametrized with constant speed. By assumption $E(\gamma) \leq E(\eta)$, so using (5.2.2)

$$L(\gamma)^2 \leq 2E(\gamma)(b-a) \leq 2E(\eta)(b-a) = L(\eta)^2.$$

Since the length of a curve does not depend on its parametrization, this shows that γ is a minimizing curve. Due to the result of (a), γ is a geodesic, up to reparametrization. Finally, we observe that γ must be parametrized with constant speed for otherwise it would not minimize the energy by the same (5.2.2) and the condition of equality thereto pertaining. $\square$

5.3 Variations of curves

A *variation* of a piecewise smooth curve $\gamma : [a, b] \to M$ is a continuous map $H : [a, b] \times (-\epsilon, \epsilon) \to M$, where $\epsilon > 0$, such that $H(s, 0) = \gamma(s)$ for all $s \in [a, b]$, and there exists a subdivision

$$a = s_0 < s_1 < \cdots < s_n = b$$

such that $H|_{[s_{i-1}, s_i]\times(-\epsilon,\epsilon)}$ is smooth for all $i = 1, \ldots, n$. For each $t \in (-\epsilon, \epsilon)$, the curve

$$t \mapsto H(s, t)$$

will be denoted by γ_t. We say that H is a *variation with fixed endpoints* if H is a variation satisfying

$$H(a, t) = \gamma_t(a) = \gamma(a) \quad \text{and} \quad H(b, t) = \gamma_t(b) = \gamma(b)$$

for every $t \in (-\epsilon, \epsilon)$. A variation H is called *smooth* if $H : [a, b] \times (-\epsilon, \epsilon) \to M$ is smooth. Finally, we say that H is a *variation through geodesics* if H is a variation such that γ_t is a geodesic for every $t \in (-\epsilon, \epsilon)$.

For a variation H of a piecewise smooth curve $\gamma : [a,b] \to M$, we will denote by $\overline{\nabla}$ the connection induced along H according to Proposition 2.6.1, and we will consider the following vector fields along H:

$$\frac{\bar\partial}{\partial t} = dH\left(\frac{\partial}{\partial t}\right) \quad \text{and} \quad \frac{\bar\partial}{\partial s} = dH\left(\frac{\partial}{\partial s}\right).$$

Note that

$$\frac{\bar\partial}{\partial s} = \gamma_t'$$

may be discontinuous at $s = s_i$. On the other hand, $\frac{\bar\partial}{\partial t}$ and $\overline{\nabla}_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial t}$ are continuous vector fields; this is true because $[a,b]\times(-\epsilon,\epsilon) = \cup_{i=1}^n [s_{i-1}, s_i]\times(-\epsilon,\epsilon)$ is a decomposition into a finite union of closed subsets, and the restrictions of those vector fields to $[s_{i-1}, s_i]\times(-\epsilon,\epsilon)$ are continuous for $i = 1, \ldots, n$. Hence we have that

$$Y = \frac{\bar\partial}{\partial t}\Big|_{t=0}$$

is a piecewise smooth vector field along γ called the *variational vector field associated to H*. Conversely, we have the following result.

5.3.1 Lemma *Given a piecewise smooth vector field Y along a piecewise smooth curve $\gamma : [a,b] \to M$, there exists a piecewise smooth variation H of γ whose associated variational vector field is Y.*

Proof. Set $H(s,t) = \exp_{\gamma(s)}(tY(s))$. Since the interval $[a,b]$ is compact, we can find $\epsilon > 0$ such that H is well defined on $[a,b]\times(-\epsilon,\epsilon)$, and

$$\frac{\bar\partial}{\partial t}\Big|_{t=0} = d(\exp_{\gamma(s)})_{0_{\gamma(s)}}(Y(s)) = Y(s).$$

□

5.3.2 Proposition (First variation of energy) *Let $\gamma : [a,b] \to M$ be a piecewise smooth curve, and let H be a variation of γ with associated variational vector field Y. Then*

$$\frac{d}{dt}\Big|_{t=0} E(\gamma_t) = \sum_{i=1}^n \langle Y, \gamma'\rangle\Big|_{s_{i-1}^+}^{s_i^-} - \int_a^b \langle Y, \overline{\nabla}_{\frac{\partial}{\partial s}}\gamma'\rangle\, ds. \tag{5.3.3}$$

Proof. Consider first the case in which γ and H are smooth. Then the integrand of

$$E(\gamma_t) = \frac{1}{2}\int_a^b \langle \gamma_t', \gamma_t'\rangle\, ds = \frac{1}{2}\int_a^b \langle \frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial s}\rangle\, ds$$

is smooth and we can compute $\frac{d}{dt}E(\gamma_t)$ by differentiation under the integral sign, namely,

$$\begin{aligned}
\frac{d}{dt}E(\gamma_t) &= \frac{1}{2}\int_a^b \frac{\partial}{\partial t}\langle \frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial s}\rangle\, ds \\
&= \int_a^b \langle \overline{\nabla}_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial s}\rangle\, ds \\
&= \int_a^b \langle \overline{\nabla}_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s}\rangle\, ds \\
&= \int_a^b \frac{\partial}{\partial s}\langle \frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s}\rangle - \langle \frac{\bar\partial}{\partial t}, \overline{\nabla}_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial s}\rangle\, ds.
\end{aligned} \tag{5.3.4}$$

Here we have used that $\overline{\nabla}_{\frac{\partial}{\partial t}}\frac{\bar{\partial}}{\partial s}-\overline{\nabla}_{\frac{\partial}{\partial s}}\frac{\bar{\partial}}{\partial t} = H_*[\frac{\partial}{\partial t},\frac{\partial}{\partial s}] = 0$, according to Proposition 2.6.2. Evaluating the above formula at $t=0$ gives the desired formula in the case in which γ and H are smooth:

$$\frac{d}{dt}\Big|_{t=0}E(\gamma_t) = \langle Y,\gamma'\rangle\Big|_{a^+}^{b^-} - \int_a^b \langle Y, \overline{\nabla}_{\frac{\partial}{\partial s}}\gamma'\rangle\, ds.$$

The formula in the general case is obtained from this one by observing that the energy is additive over a union of subintervals. ☐

5.3.5 Proposition (Critical points of E) *Let $\gamma:[a,b]\to M$ be a piecewise smooth curve. We have that*

$$\frac{d}{dt}\Big|_{t=0}E(\gamma_t) = 0$$

for every variation with fixed endpoints if and only if γ is a geodesic.

Proof. In the class of variations with fixed endpoints, we have that $Y(a)=Y(b)=0$, so formula (5.3.3) can be rewritten as

$$\frac{d}{dt}\Big|_{t=0}E(\gamma_t) = -\sum_{i=1}^{n-1}\langle Y,\gamma'\rangle\Big|_{s_i^-}^{s_i^+} - \int_a^b \langle Y, \overline{\nabla}_{\frac{\partial}{\partial s}}\gamma'\rangle\, ds. \tag{5.3.6}$$

If γ is a geodesic, then $\bar{\nabla}_{\frac{\partial}{\partial s}}\gamma'=0$ and γ' is continuous, so both terms in (5.3.6) vanish proving one direction of the proposition.

Conversely, suppose that $0=\frac{d}{dt}\big|_{t=0}E(\gamma_t)=0$ for every variation with fixed endpoints. Let $f:[a,b]\to\mathbf{R}$ be a smooth function such that $f(s)>0$ if $s\neq s_i$ and $f(s_i)=0$ for $i=0,\ldots,n$, and set $Y=f\bar{\nabla}_{\frac{\partial}{\partial s}}\gamma'$. Then Y is a piecewise smooth vector field along γ (note that Y is indeed continuous at s_i) with $Y(a)=Y(b)=0$, and so it defines via Lemma 5.3.1 a variation $\{\gamma_t\}$ with fixed endpoints for which (5.3.6) gives that $0=-\int_a^b f||\overline{\nabla}_{\frac{\partial}{\partial s}}\gamma'||^2\,ds$. This already implies that γ is a geodesic on (s_{i-1},s_i) for $i=1,\ldots,n$. Since $\gamma|_{[s_{i-1},s_i]}$ is smooth by assumption, it follows that $\overline{\nabla}_{\frac{\partial}{\partial s}}\gamma'|_{s_i}=0$ in the sense of side derivatives.

Next, we take Y to be a smooth vector field along γ satisfying $Y(a)=Y(b)=0$ and $Y(s_i)=\gamma'(s_i^+)-\gamma'(s_i^-)$ for $i=2,\ldots,n-1$. Substituting into (5.3.6) now gives that $0=-\sum_{i=2}^{n-1}||\gamma'(s_i^+)-\gamma'(s_i^-)||^2$. This of course implies that γ is of class C^1. Since we already know that $\gamma|_{[s_{i-1},s_i]}$ is a geodesic for $i=1,\ldots,n$, this implies that these restrictions are segments of the same geodesic γ defined on $[a,b]$ by the uniqueness result (Proposition 2.4.3). ☐

5.3.7 Corollary (Critical points of L) *Let $\gamma:[a,b]\to M$ be a piecewise smooth curve. We have that*

$$\frac{d}{dt}\Big|_{t=0}L(\gamma_t) = 0$$

for every variation with fixed endpoints if and only if γ is a geodesic, up to reparametrization.

Proof. Let $\tilde{\gamma}=\gamma\circ\varphi$ be a reparametrization of γ with constant speed, where $\varphi:[a,b]\to[a,b]$ is an orientation-preserving diffeomorphism. Given a variation H with fixed endpoints of γ, we define a variation $\tilde{H}$ of $\tilde{\gamma}$ by setting $\tilde{H}(s,t)=H(\varphi(s),t)$, and we denote $\tilde{\gamma}_t(s)=\tilde{H}(s,t)=(\gamma_t\circ\varphi)(s)$. Of course $L(\gamma_t)=L(\tilde{\gamma}_t)$, so we may assume without loss of generality that γ is parametrized with constant speed from the outset. Now

$$\frac{d}{dt}L(\gamma_t) = \int_a^b \frac{\partial}{\partial t}\langle\frac{\bar{\partial}}{\partial s},\frac{\bar{\partial}}{\partial s}\rangle^{1/2}\,ds = \frac{1}{2}\int_a^b \langle\frac{\bar{\partial}}{\partial s},\frac{\bar{\partial}}{\partial s}\rangle^{-1/2}\frac{\partial}{\partial t}\langle\frac{\bar{\partial}}{\partial s},\frac{\bar{\partial}}{\partial s}\rangle\,ds.$$

Evaluating at $t = 0$ and using that $||\gamma'||$ is a constant $k \neq 0$ gives that

$$\frac{d}{dt}\Big|_{t=0} L(\gamma_t) = \frac{1}{2k}\int_a^b \frac{\partial}{\partial t}\Big|_{t=0}\langle \frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial s}\rangle\, ds = \frac{1}{k}\frac{d}{dt}\Big|_{t=0} E(\gamma_t).$$

This shows that L and E have the same critical points, up to reparametrization. Thus the desired result is an immediate consequence of Proposition 5.3.5. □

5.3.8 Proposition (Second variation of energy) *Let $\gamma : [a, b] \to M$ be a geodesic, and let H be a piecewise smooth variation of γ with associated variational vector field Y. Then*

$$\frac{d^2}{dt^2}\Big|_{t=0} E(\gamma_t) = \langle \overline\nabla_{\frac{\partial}{\partial t}} \frac{\bar\partial}{\partial t}\Big|_{t=0}, \gamma'\rangle\Big|_a^b + \int_a^b ||Y'||^2 + \langle R(\gamma', Y)\gamma', Y\rangle\, ds, \tag{5.3.9}$$

where $Y' = \frac{\nabla Y}{ds}$.

Proof. Starting with formula (5.3.4), we compute that

$$\begin{aligned}
\frac{d^2}{dt^2}E(\gamma_t) &= \int_a^b \frac{\partial}{\partial t}\langle \overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial s}\rangle\, ds \\
&= \int_a^b \frac{\partial}{\partial t}\langle \overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s}\rangle\, ds \\
&= \int_a^b \langle \overline\nabla_{\frac{\partial}{\partial t}}\overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s}\rangle + \langle \overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial t}, \overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial s}\rangle\, ds \\
&= \int_a^b \langle \overline\nabla_{\frac{\partial}{\partial s}}\overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s}\rangle + \langle R(\frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s})\frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s}\rangle + \Big|\Big|\overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial t}\Big|\Big|^2 ds \\
&= \int_a^b \frac{\partial}{\partial s}\langle \overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s}\rangle - \langle \overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial t}, \overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial s}\rangle + \langle R(\frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial t})\frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial t}\rangle + \Big|\Big|\overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial t}\Big|\Big|^2 ds
\end{aligned}$$

In the fourth equality, we used that $\overline\nabla_{\frac{\partial}{\partial t}}\overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial t} - \overline\nabla_{\frac{\partial}{\partial s}}\overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial t} = R(\frac{\bar\partial}{\partial t}, \frac{\bar\partial}{\partial s})\frac{\bar\partial}{\partial t}$, according to exercise 11 of chapter 4. Evaluating this formula at $t = 0$ yields that

$$\frac{d^2}{dt^2}\Big|_{t=0} E(\gamma_t) = \int_a^b \frac{\partial}{\partial s}\langle \overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial t}\Big|_{t=0}, \gamma'\rangle - \langle \overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial t}\Big|_{t=0}, \gamma''\rangle + \langle R(\gamma', Y)\gamma', Y\rangle + ||Y'||^2\, ds$$

Since γ' and $\overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial t}$ are continuous and $\gamma'' = 0$, this proves the desired formula. □

5.4 Jacobi fields

Throughout this section, we fix a geodesic $\gamma : [0, \ell] \to M$. The second variation formula (5.3.9) defines a quadratic form on the space of piecewise smooth vector fields along γ vanishing at 0 and ℓ whose associated symmetric bilinear form I is called the *index form* and is clearly given by

$$I(X, Y) = \int_0^\ell \langle X', Y'\rangle + \langle R(\gamma', X)\gamma', Y\rangle\, ds,$$

where $X' = \frac{\nabla X}{ds}$, $Y' = \frac{\nabla Y}{ds}$. Let $0 = s_0 < s_1 < \cdots < s_n = \ell$ be a subdivision of $[0, \ell]$ such that X and Y are smooth on $[s_{i-1}, s_i]$ for $i = 1, \ldots, n$. Since $\langle X', Y'\rangle = \langle X, Y'\rangle' - \langle X, Y''\rangle$ on each

$[s_{i-1}, s_i]$, we can write

$$
\begin{aligned}
I(X,Y) &= \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} \langle X, Y' \rangle' \, ds + \int_0^\ell -\langle X, Y'' \rangle + \langle R(\gamma', Y)\gamma', X \rangle \, ds \\
&= \sum_{i=1}^{n} \langle X, Y' \rangle \Big|_{s_{i-1}^+}^{s_i^-} + \int_0^\ell \langle -Y'' + R(\gamma', Y)\gamma', X \rangle \, ds \\
&= -\sum_{i=1}^{n-1} \langle Y'(s_i^+) - Y'(s_i^-), X \rangle + \int_0^\ell \langle -Y'' + R(\gamma', Y)\gamma', X \rangle \, ds \qquad (5.4.1)
\end{aligned}
$$

A *Jacobi field along* γ is a smooth vector field Y along γ (not necessarily vanishing at the endpoints of γ) such that

$$
-Y'' + R(\gamma', Y)\gamma' = 0. \qquad (5.4.2)
$$

Hence the space of Jacobi fields along γ vanishing at the endpoints of γ is contained in the kernel of I as a bilinear form; it is easy to show that these spaces in fact coincide by using ideas very similar to the ones in the proof of Proposition 5.3.5 (cf. exercise 2). Equation (5.4.2) is called the *Jacobi equation along* γ.

Next, denote by $\mathcal{J}$ the space of all Jacobi fields along γ. It is obvious that $\mathcal{J}$ is a vector space. It is also a very simple matter to check that the smooth vector fields along γ given by $Y_0(s) = \gamma'(s)$ and $Y_1(s) = s\gamma'(s)$ belong to $\mathcal{J}$. The next proposition shows that a Jacobi field Y along γ, being a solution of a second-order linear ordinary differential equation, is completely determined by its initial conditions $Y(0) \in T_pM$ and $Y'(0) \in T_pM$. It follows that $\mathcal{J}$ is a finite-dimensional vector space and $\dim \mathcal{J} = 2 \dim M$.

5.4.3 Proposition *Let $\gamma : [0, \ell] \to M$ be a geodesic, and put $\gamma(0) = p$.*

a. Given u, $v \in T_pM$, there exists a unique Jacobi field $Y \in \mathcal{J}$ such that $Y(0) = u$ and $Y'(0) = v$.

b. If X, $Y \in \mathcal{J}$, then the function $\langle X', Y \rangle - \langle X, Y' \rangle$ is constant on $[0, \ell]$. It follows that $\langle \gamma'(s), Y(s) \rangle = as + b$ for some constants a, $b \in \mathbf{R}$ and $s \in [0, \ell]$.

Proof. (a) Select an orthonormal basis $\{e_1, \ldots, e_n\}$ of T_pM with $e_1 = \gamma'(0)$ and extend it to an orthonormal frame $\{E_1, \ldots, E_n\}$ of parallel vector fields along γ; since γ is a geodesic, $E_1 = \gamma'$. Let Y be a smooth vector field along γ. Then we can write $Y = \sum_{i=1}^n f_i E_i$, where $f_i : [0, \ell] \to \mathbf{R}$ are smooth functions. In these terms, the Jacobi equation (5.4.2) is

$$
\sum_{i=1}^{n} -f_i'' E_i + f_i R(\gamma', E_i)\gamma' = 0.
$$

Taking the inner product of the left-hand side with E_j yields that

$$
-f_j'' + \sum_{i=2}^{n} \langle R(\gamma', E_i)\gamma', E_j \rangle f_i = 0
$$

for $j = 1, \ldots, n$. This is a system of second-order ordinary linear differential equations for which the standard theorems of existence and uniqueness of solutions apply, hence the result.

(b) In order to prove the constancy of the function, it suffices to differentiate it along γ:

$$\begin{aligned}(\langle X', Y\rangle - \langle X, Y'\rangle)' &= (\langle X'', Y\rangle + \langle X', Y'\rangle) - (\langle X', Y'\rangle + \langle X, Y''\rangle) \\ &= \langle R(\gamma', X)\gamma', Y\rangle - \langle X, R(\gamma', Y)\gamma'\rangle \\ &= 0,\end{aligned}$$

where we have used the Jacobi equation (5.4.2) and the symmetry of R (Proposition 4.2.1(c)).

Finally, in order to get the last assertion, take $X = \gamma'$ in the function. Then $\langle \gamma', Y'\rangle = \langle \gamma', Y\rangle'$ is a constant. It follows that $\langle \gamma', Y\rangle$ has the required form. □

Proposition 5.4.3(b) shows that $Y \in \mathcal{J}$ satisfies $\langle \gamma'(s), Y(s)\rangle = as + b$ for all $s \in [0, \ell]$ where $a = \langle \gamma'(0), Y'(0)\rangle$ and $b = \langle \gamma'(0), Y(0)\rangle$. Writing

$$Y = (Y - aY_1 - bY_0) + bY_0 + aY_1$$

shows that there exists a splitting

$$\mathcal{J} = \mathcal{J}^\perp \oplus \mathbf{R}Y_0 \oplus \mathbf{R}Y_1,$$

where $\mathcal{J}^\perp$ is the subspace of Jacobi fields along γ that are always orthogonal to γ', namely,

$$\mathcal{J}^\perp = \{\, Y \in \mathcal{J} \mid \langle Y(s), \gamma'(s)\rangle = 0 \quad \text{for all } s \in [0, \ell] \,\}.$$

Since Y_0 and Y_1 *always* belong to $\mathcal{J}$, it is the subspace $\mathcal{J}^\perp$ that can give us effective information about the geodesic γ, if any.

The next proposition refines the information of Lemma 5.3.1. It also points out the fact that the Jacobi fields along a geodesic somehow control the behaviour of the nearby geodesics.

5.4.4 Proposition *Let $\gamma : [0, \ell] \to M$ be a geodesic. If H is a smooth variation of γ through geodesics, then the associated variational vector field Y is a Jacobi field along γ. On the other hand, every Jacobi field Y along γ is the variational vector field associated to a variation H of γ through geodesics.*

Proof. Suppose first that H is a smooth variation of γ through geodesics and let $Y = \frac{\bar\partial}{\partial t}\big|_{t=0}$ be the associated variational vector field. Then, $\overline\nabla_{\frac{\partial}{\partial s}} \frac{\bar\partial}{\partial s} = 0$, so using exercise 11 of chapter 4,

$$\overline\nabla_{\frac{\partial}{\partial s}}\overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial t} = \overline\nabla_{\frac{\partial}{\partial s}}\overline\nabla_{\frac{\partial}{\partial t}}\frac{\bar\partial}{\partial s} = \overline\nabla_{\frac{\partial}{\partial t}}\overline\nabla_{\frac{\partial}{\partial s}}\frac{\bar\partial}{\partial s} + R(\frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial t})\frac{\bar\partial}{\partial s} = R(\frac{\bar\partial}{\partial s}, \frac{\bar\partial}{\partial t})\frac{\bar\partial}{\partial s}.$$

Evaluating this formula at $t = 0$ gives that $Y'' = R(\gamma', Y)\gamma'$, and hence, Y is a Jacobi field.

Suppose now that Y is a Jacobi field along γ. We construct a variation H of γ as follows. Take any smooth curve η satisfying $\eta(0) = \gamma(0)$ and $\eta'(0) = Y(0)$. Let X_0 and X_1 be the parallel vector fields along η such that $X_0(0) = \gamma'(0)$ and $X_1(0) = Y'(0)$, and let $X(t) = X_0(t) + tX_1(t)$. Finally, set $H(s, t) = \exp_{\eta(t)}(sX(t))$.

By construction, H is a variation through geodesics, so $\frac{\bar\partial}{\partial t}\big|_{t=0} = dH(\frac{\partial}{\partial t})\big|_{t=0}$ is a Jacobi field along γ by the first part of this proof. Let us compute the initial conditions of $\frac{\bar\partial}{\partial t}\big|_{t=0}$ at $s = 0$. Since $H(0, t) = \eta(t)$, we have

$$\frac{\bar\partial}{\partial t}\Big|_{\substack{t=0\\ s=0}} = \eta'(0) = Y(0).$$

Moreover,

$$\frac{\bar\partial}{\partial s}\Big|_{s=0} = d(\exp_{\eta(t)})_{0_{\eta(t)}}(X(t)) = X(t),$$

so

$$\bar\nabla_{\frac{\partial}{\partial s}} \frac{\bar\partial}{\partial t}\Big|_{\substack{t=0\\ s=0}} = \bar\nabla_{\frac{\partial}{\partial t}} \frac{\bar\partial}{\partial s}\Big|_{\substack{t=0\\ s=0}} = X'(0) = X_1(0) = Y'(0).$$

Since $\frac{\bar\partial}{\partial t}\big|_{t=0}$ and Y are Jacobi fields along γ having the same initial conditions at $s = 0$, they are equal, and this finishes the proof of the proposition. □

5.4.5 Scholium *Consider a point $p \in M$ and two tangent vectors $u, v \in T_pM$. Let γ be the geodesic $\gamma(s) = \exp_p(sv)$, and let Y be the Jacobi field along γ satisfying $Y(0) = 0$ and $Y'(0) = u$. Then*

$$Y(s) = d(\exp_p)_{sv}(su)$$

for all s in the domain of γ.

Proof. This proof is contained in the proof of second assertion in the statement of Proposition 5.4.4. Indeed, using the notation from that proof, η is the constant curve at p, X_0 is the constant vector field $\gamma'(0) = v$ and X_1 is the constant vector field $Y'(0) = u$, so $H(s,t) = \exp_p(s(v+tu))$ and

$$Y(s) = \frac{\bar\partial}{\partial t}\Big|_{(s,0)} = d(\exp_p)_{sv}(su),$$

as desired. □

5.4.6 Example In special cases, knowledge of the Jacobi fields can be used to compute the sectional curvature. Recall the surface of revolution in $\mathbf{R}^3$ as in Example 1.2.2(b). Note that the meridians $\theta =$ const. are geodesics by the reflection argument used in the case of S^n (cf. page 57). By rotational symmetry, it suffices to compute the sectional curvature along the meridian $\gamma(s) = \varphi(s,0)$. We produce a variation of γ by using nearby meridians, namely $H(s,t) = \mathbf{x}(s,t)$. In this case the Jacobi field is $Y(s) = \frac{\bar\partial}{\partial t}\big|_{(s,0)} = \mathbf{x}_\theta(s,0) = f(s)\frac{\partial}{\partial y}$. Note that $\{\gamma', \frac{\partial}{\partial y}\}$ is a parallel orthonormal frame along γ. Therefore the Jacobi equation (5.4.2) is $-f''(s) - K(s)f(s) = 0$, where K is the Gaussian curvature along the parallel $\mathbf{x}(s,\cdot)$. Hence $K = -f''/f$.

5.5 Conjugate points

Let $\gamma(s) = \exp_p(sv)$ be a geodesic in M, where $p \in M$ and $v \in T_pM$. A point $\gamma(s_0)$, where $s_0 > 0$, is called a *point conjugate to p along γ* or a *conjugate point of p along γ* if there exists a nontrivial Jacobi field Y along γ such that $Y(0) = 0$ and $Y(s_0) = 0$; the parameter value s_0 is called a *conjugate value*. In this case, we also have that p is conjugate to $\gamma(s_0)$ along γ^{-1}, so we sometimes say that p and $\gamma(s_0)$ are *conjugate points along γ*. A point $q \in M$ is called a *point conjugate to p* if q is conjugate to p along some geodesic emanating from p. The set of all points of M conjugate to p is called the *conjugate locus of p*.

If $q = \gamma(s_0)$ is conjugate to p along $\gamma(s) = \exp_p(sv)$, and Y is a Jacobi field along γ such that $Y(0) = 0$ and $Y(s_0) = 0$, then Y is everywhere perpendicular to γ' by Proposition 5.4.3(b). Even more interesting, $Y'(0)$ lies in the kernel of the map $d(\exp_p)_{s_0v}$ as it follows from Scholium 5.4.5. Hence, the points conjugate to p are exactly the critical values of $\exp_p$. The *multiplicity* of q as a point conjugate to p along γ is the dimension of the kernel of $d(\exp_p)_{s_0v}$.

Intuitively speaking, the meaning of q being a conjugate point of p along a geodesic γ is that some nearby geodesics emanating from p must meet γ at q *at least in the infinitesimal sense*. Before proceeding with the main result of this section, we prove two lemmas.

5.5.1 Lemma (Gauss, global version) *Consider a point $p \in M$, two tangent vectors $u, v \in T_pM$, and the geodesic $\gamma(s) = \exp_p(sv)$. Then*

$$g_{\gamma(s)}\big(\, d(\exp_p)_{sv}(u), d(\exp_p)_{sv}(v)\,\big) = g_p(u,v).$$

Proof. Note the right-hand-side in the formula is the value at $s = 0$ of the left-hand-side of it. Note also that $d(\exp_p)_{sv}(v) = \gamma'(s)$. Next, let Y denote the Jacobi field along γ with initial conditions $Y(0) = 0$ and $Y'(0) = u$. On the one hand, we know from Scholium 5.4.5 that $d(\exp_p)_{sv}(u) = \frac{1}{s}Y(s)$ for $s \neq 0$. On the other hand, decompose $u = \lambda v + u_1$, where u_1 is perpendicular to v, and let Y_0, Y_1 be the Jacobi fields along γ vanishing at $s = 0$ such that $Y_0'(0) = \lambda v$ and $Y_1'(0) = u_1$. Then $Y_0(s) = \lambda s\gamma'(s)$ and $Y(s) = Y_0(s) + Y_1(s) = \lambda s\gamma'(s) + Y_1(s)$, so, if $s \neq 0$,

$$\begin{aligned} g_{\gamma(s)}\big(\, d(\exp_p)_{sv}(u), d(\exp_p)_{sv}(v)\,\big) &= g_{\gamma(s)}\big(\,\frac{1}{s}Y(s), \gamma'(s)\,\big) \\ &= \lambda g_{\gamma(s)}\big(\,\gamma'(s), \gamma'(s)\,\big) + \frac{1}{s} g_{\gamma(s)}\big(\,Y_1(s), \gamma'(s)\,\big). \end{aligned}$$

The first term in the last line of the above calculation is $\lambda g_p(v,v) = g_p(u,v)$, since the length of the tangent vector of a geodesic is constant. The second term in there is zero by Proposition 5.4.3(b) because $Y_1(0)$ and $Y_1'(0)$ are perpendicular to $\gamma'(0)$, and this proves the formula. □

5.5.2 Lemma *Consider a point $p \in M$, and a tangent vector $v \in T_pM$. Let $\varphi : [0,1] \to T_pM$ denote the radial segment $\varphi(s) = sv$, and let $\psi : [0,1] \to T_pM$ be an arbitrary piecewise smooth curve joining the origin* 0 *to v. Then*

$$L(\exp_p \circ\psi) \geq L(\exp_p \circ\varphi) = ||v||.$$

Proof. Without loss of generality, we may assume that $\psi(s) \neq 0$ for $s > 0$. In the case in which ψ is smooth, write $\psi(s) = r(s)u(s)$ where $r : (0,1] \to (0,+\infty)$ and $u : (0,1] \to S^{n-1}$ are smooth, and S^{n-1} denotes the unit sphere of (T_pM, g_p). Then

$$\psi'(s) = r'(s)u(s) + r(s)u'(s)$$

with $\langle u(s), u'(s)\rangle = 0$. Applying Gauss lemma 5.5.1 twice in the following computation,

$$\begin{aligned} ||(\exp_p \circ\psi)'(s)||^2 &= ||d(\exp_p)_{\psi(s)}(\psi'(s))||^2 \\ &= (r'(s))^2 \underbrace{||d(\exp_p)_{\psi(s)}(u(s))||^2}_{=||u(s)||^2=1} + (r(s))^2||d(\exp_p)_{\psi(s)}(u'(s))||^2 \\ &\geq (r'(s))^2, \end{aligned}$$

we get that

$$L(\exp_p \circ\psi) \geq \int_0^1 |r'(s)|\, ds \geq |r(1) - \lim_{s\to 0+} r(s)| = ||v||.$$

In the general case, we repeat the argument above over each subinterval where ψ is smooth and add up the estimates. □

Next, we prove the main result of this chapter. It gives a sufficient condition and a necessary condition for a geodesic segment to be locally minimizing is the space of curves with the same endpoints.

5.5.3 Theorem (Jacobi-Darboux) *Let $\gamma : [0, \ell] \to M$ be a geodesic segment parametrized with unit speed and with endpoints $\gamma(0) = p$ and $\gamma(\ell) = q$.*

a. If there are no points conjugate to p along γ, then there exists a neighborhood V of γ in the C^0-topology (or uniform topology) in the space of piecewise smooth curves parametrized on $[0, \ell]$ and joining p to q such that $E(\eta) \geq E(\gamma)$ and $L(\eta) \geq L(\gamma)$ for every $\eta \in V$. Moreover, if $L(\eta) = L(\gamma)$ for some $\eta \in V$, then η and γ differ by a reparametrization.

b. If $\gamma(s_0)$ is conjugate to p along γ for some $s_0 \in (0, \ell)$, then there exists a variation $\{\gamma_t\}$ of γ with fixed endpoints such that $E(\gamma_t) < E(\gamma)$ and $L(\gamma_t) < L(\gamma)$ for sufficiently small t.

Proof. Put $\gamma'(0) = v$ and define $\varphi : [0, \ell] \to T_pM$ by $\varphi(s) = sv$. By assumption, $\varphi(s)$ is a regular point of $\exp_p$ for $s \in [0, \ell]$. Since $\varphi([0, \ell])$ is compact, we can cover it by a union $\cup_{i=1}^k W_i$ of open balls $W_i \subset T_pM$ such that $\exp_p$ is a diffeomorphism of W_i onto an open subset $U_i \subset M$. Choose a subdivision $0 = s_0 < s_1 < \ldots < s_k = \ell$ such that $\varphi([s_{i-1}, s_i]) \subset W_i$ for all i. Let V be the open ball centered at γ of radius $\epsilon > 0$, namely, V consists of the piecewise smooth curves $\eta : [0, \ell] \to M$ joining p to q and satisfying $d(\eta(s), \gamma(s)) < \epsilon$ for $s \in [0, \ell]$. We take ϵ so that $\eta([s_{i-1}, s_i]) \subset U_i$ for $\eta \in V$ and $i = 1, \ldots, k$. Note that $\exp_p(W_{i-1} \cap W_i)$ is an open neighborhood of $\gamma(s_{i-1})$ contained in $U_{i-1} \cap U_i$. We further decrease ϵ, if necessary, so as to obtain that $\eta(s_{i-1}) \in \exp_p(W_{i-1} \cap W_i)$ for $\eta \in V$ and $i = 2, \ldots, k$.

For each $\eta \in V$, we lift η to a piecewise smooth curve ψ in T_pM as follows. Define

$$\psi(s) = (\exp_p|_{W_1})^{-1}(\eta(s)) \quad \text{for } s \in [0, s_1].$$

Note that $\psi(0) = 0$. Assume that ψ has already been defined on $[0, s_{i-1}]$ for some $2 \leq i \leq k$ such that it satisfies $\exp_p(\psi(s)) = \eta(s)$ for $s \in [0, s_{i-1}]$ and $\psi(s_{i-1}) \in W_{i-1}$. Note that these conditions imply that

$$\exp_p(\psi(s_{i-1})) = \eta(s_{i-1}) \in \exp_p(W_{i-1} \cap W_i),$$

so $\psi(s_{i-1}) \in W_i$. Hence

$$\psi(s) = (\exp_p|_{W_i})^{-1}(\eta(s)) \quad \text{for } s \in [s_{i-1}, s_i]$$

continuously extends ψ to $[0, s_i]$. This completes the induction step and shows that ψ can be defined on $[0, \ell]$. Since $\eta(\ell) \in W_k$, we have $\psi(\ell) = \ell v$. By Lemma 5.5.2,

$$L(\eta) = L(\exp_p \circ \psi) \geq L(\exp_p \circ \varphi) = L(\gamma).$$

Moreover, since $d(\exp_p)_{\psi(s)}$ is injective for $s \in [0, \ell]$, the proof of the lemma shows that the inequality is sharp unless u is constant and r' is nonnegative in the notation of that proof, that is, η coincides with γ up to reparametrization. As for the assertion concerning the energy, we observe that

$$E(\eta) \geq \frac{1}{2\ell} L(\eta)^2 \geq \frac{1}{2\ell} L(\gamma)^2 = E(\gamma)$$

by the Cauchy-Schwarz inequality (5.2.2). This proves part (a).

(b) By assumption, there exists a nontrivial Jacobi field Y along γ such that $Y(0) = Y(s_0) = 0$. Owing to the non-triviality of Y, $Y'(s_0) \neq 0$. Let Z_1 be the parallel vector field along γ with $Z_1(s_0) = -Y'(s_0)$, construct a smooth function $\theta : [0, \ell] \to \mathbf{R}$ such that $\theta(0) = \theta(\ell) = 0$ and $\theta(s_0) = 1$, and set $Z(s) = \theta(s) Z_1(s)$. Also, extend Y to a piecewise smooth vector field on $[0, \ell]$ by putting $Y|_{[s_0, \ell]} = 0$, and set $Y_\alpha(s) = Y(s) + \alpha Z(s)$ for $s \in [0, \ell]$ and $\alpha \in \mathbf{R}$.

Now Y_α is a piecewise smooth vector field along γ which is everywhere normal to γ' and vanishes at 0 and ℓ. Consider a variation with fixed endpoints $\{\gamma_t\}$ with associated variational vector field Y_α. Then

$$\begin{aligned} I(Y_\alpha, Y_\alpha) &= I(Y,Y) + 2\alpha I(Y,Z) + \alpha^2 I(Z,Z) \\ &= -2\alpha\langle Y'(s_0^+) - Y'(s_0^-), Z(s_0)\rangle + \alpha^2 I(Z,Z) \\ &= -2\alpha||Y'(s_0^-)||^2 + \alpha^2 I(Z,Z) \\ &< 0, \end{aligned}$$

where $\alpha > 0$ is chosen sufficiently small so as to ensure the last inequality. Hence $E(\gamma_t) < E(\gamma)$ for sufficiently small t. Also,

$$L(\gamma_t)^2 \leq 2\ell E(\gamma_t) < 2\ell E(\gamma) = L(\gamma)^2,$$

and this completes the proof. □

As a corollary of the theorem of Jacobi-Darboux 5.5.3, we have the following refinement of Proposition 3.4.3.

5.5.4 Corollary *Let M be a complete Riemannian manifold. Then, for each $p \in M$, the exponential map*

$$\exp_p : \mathrm{D}_p \to M \setminus \mathrm{Cut}(p)$$

is a diffeomorphism.

Proof. We have already seen that $\exp_p(\mathrm{D}_p) = M\setminus\mathrm{Cut}(p)$. Theorem 5.5.3 implies that a geodesic $\gamma_v : [0, +\infty) \to M$, where $v \in T_pM$ and $||v|| = 1$, does not minimize L past its first conjugate point, so a conjugate point along γ_v, if existing, must occur at a parameter value $s_0 \geq \rho(v)$. It follows that $\exp_p$ is a local diffeomorphism at sv for $s \in [0, \rho(v))$. Since v is an arbitrary unit tangent vector at p, this shows that $\exp_p$ is a local diffeomorphism on D_p. It remains only to check that $\exp_p$ is injective on D_p. But this is clear since any point in $\exp_p(\mathrm{D}_p)$ can be joined to p by a unique minimal geodesic as was already observed right after the proof of Proposition 3.4.3. □

The *first conjugate point* along a geodesic $\gamma(s) = \exp_p(sv)$, where $p \in M$ and $v \in T_pM$, is the smallest parameter value $s_0 > 0$ such that $\gamma(s_0)$ is conjugate to p along γ. It also follows from the theorem of Jacobi-Darboux 5.5.3 that the first conjugate point to p along γ cannot occur before the cut point; in particular, the conjugate locus of a point is empty if its cut locus is empty. The following proposition gives more information.

5.5.5 Proposition *Let M be a complete Riemannian manifold, and let $p \in M$. Then a point q belongs to the cut locus* $\mathrm{Cut}(p)$ *if and only if one of the following non-mutually exclusive assertions is true:*

a. There exist at least two distinct minimizing geodesics joining p to q.

b. The point q is the first conjugate point to p along a minimizing geodesic.

In particular, $q \in \mathrm{Cut}(p)$ if and only if $p \in \mathrm{Cut}(q)$.

Proof. By Lemma 3.4.1 and Theorem 5.5.3, we already know that the conditions in the statement are sufficient for q to belong to $\mathrm{Cut}(p)$. Conversely, suppose that $q \in \mathrm{Cut}(p)$. Then we can write $q = \exp_p(\rho(v)v)$ for some unit vector $v \in T_pM$ with $\rho(v) < +\infty$. In particular, $\gamma(s) = \exp_p(sv)$, where $0 \leq s \leq \rho(v)$, is a minimal geodesic joining p to q. Choose a sequence (s_j) of real numbers such that $s_j \searrow \rho(v)$. For each j, there exists a minimal geodesic γ_j joining p to $\gamma(s_j)$, say $\gamma_j(s) =$

$\exp_p(sw_j)$, where $w_j \in T_pM$ and $||w_j|| = 1$. Let $d_j = d(p, \gamma(s_j))$, so that $\gamma_j(d_j) = \gamma(s_j)$. Since $s_j > \rho(v)$, we have that $\gamma|_{[0,s_j]}$ is not minimal so that $d_j < s_j$.

Next, by compactness of the unit sphere in T_pM and by passing to a subsequence if necessary, we may assume that (w_j) converges to a unit vector $w \in T_pM$. Since the distance d is continuous, $d_j = d(p, \gamma(s_j)) \to d(p, \gamma(\rho(v))) = \rho(v)$. By taking the limit as $j \to +\infty$ in $\gamma(s_j) = \gamma_j(d_j) = \exp_p(d_jw_j)$, we get that $q = \exp_p(\rho(v)w)$. Now there are two cases to be considered.

If $w \neq v$, then $\eta(s) = \exp_p(sw)$ is a minimizing geodesic joining p to q and $\eta \neq \gamma$, so we are in situation (a). On the other hand, if $w = v$, then we already have that $\exp_p(d_jw_j) = \gamma(s_j) = \exp_p(s_jv)$ for all j, where $d_jw_j \to \rho(v)v$ and $s_jv \to \rho(v)v$. It follows that $\exp_p$ is not locally injective at $\rho(v)v$, so $\rho(v)v$ is a singular point of $\exp_p$. Hence $q = \exp_p(\rho(v)v)$ is conjugate to p along γ. Since γ is minimizing on $[0, \rho(v)]$, q must be the first conjugate point to p along γ, and we are in situation (b).

For the last assertion, one needs to note that conditions (a) and (b) are symmetric in p and q. This is clear for (a) and follows from Theorem 5.5.3(b) for (b). □

All possibilities given by Proposition 5.5.5 for a point $q \in \mathrm{Cut}(p)$ can indeed occur: both (a) and (b); (a) and not (b); (b) and not (a). Comparing the examples in the sequel with the examples of section 3.5, one immediately finds situations in which the first two possibilities occur. However, the third possibility — in which q is the first conjugate point along a minimizing geodesic γ and there is no other minimizing geodesic from p to q — is not so easy to detect. The Heisenberg group (consisting of upper triangular real matrices of size 3 with 1's along the diagonal) equipped with some left-invariant metric provides such an example [Wal97, p. 352].

5.6 Examples

Flat manifolds

For a flat manifold, $R \equiv 0$, so the Jacobi equation is $Y'' = 0$. Hence Jacobi fields along a geodesic γ have the form $Y(s) = sE_1(s) + E_2(s)$, where E_1 and E_2 are parallel vector fields along γ. For instance, a Jacobi field Y along a geodesic γ in Euclidean space $\mathbf{R}^n$ is of the form $Y(s) = u + sv$, where u, $v \in \mathbf{R}^n$. If T^n is a flat torus and $\pi : \mathbf{R}^n \to T^n$ denotes the corresponding Riemannian covering, then a Jacobi field along the geodesic $\pi \circ \gamma$ in T^n is of the form $\bar{Y}(s) = d\pi_{\gamma(s)}(Y(s)) = d\pi_{\gamma(s)}(u) + sd\pi_{\gamma(s)}(v)$.

In particular, in a flat manifold there are no conjugate points, so any geodesic segment is a local minimum for L. Note that in a flat torus there are infinitely many geodesics with given endpoints p and q, and generically (meaning the case in which $q \notin \mathrm{Cut}(p)$) only one of them is a global minimum.

Manifolds of nonzero constant curvature

Consider first the unit sphere S^n. If γ is a unit speed geodesic and Y is a Jacobi field along γ which is everywhere perpendicular to γ', then formula (4.5.2) says that $R(\gamma', Y)\gamma' = -Y$, so the Jacobi equation is $Y'' = -Y$. It follows that $Y(s) = \cos sE_1(s) + \sin sE_2(s)$, where E_1 and E_2 are parallel vector fields along γ which are perpendicular to γ' (Note that a parallel vector field along γ which is perpendicular to γ' is nothing but a constant vector field on the surrounding $\mathbf{R}^{n+1}$ which is perpendicular to the 2-plane spanned by $\gamma(0)$ and $\gamma'(0)$.) In particular, if Y vanishes at $s = 0$, then $E_1 = 0$. Assuming Y is nontrivial, that is, $E_2 \neq 0$, then the conjugate values are $s = \pi, 2\pi, 3\pi, \ldots$. Therefore the first conjugate point of $p = \gamma(0)$ along γ is $-p$, so that the first conjugate locus coincides with the cut locus; since $Y'(0)$ can be any vector perpendicular to $\gamma'(0)$, the multiplicity of $-p$ is $n-1$. Note also that p is conjugate to itself along γ.

Consider now $\mathbf{R}P^n$. Since it has the same curvature tensor as S^n, it has also the same Jacobi equation, the same Jacobi fields and the same conjugate values. However, the difference to S^n is that now the first conjugate point $\gamma(\pi)$ along a geodesic γ coincides with $\gamma(0)$, so the first conjugate point occurs after the cut point $\gamma(\frac{\pi}{2})$. In particular, a geodesic of length $\frac{\pi}{2}+\epsilon$, $\epsilon>0$ small, is a local minimum for L, but not a global one.

The case of $\mathbf{R}H^n$ is similar to that of S^n. By (4.5.3), the Jacobi equation is $Y''=Y$, so the Jacobi fields along a geodesic γ have the form $Y(s)=\cosh sE_1(s)+\sinh sE_2(s)$, where E_1 and E_2 are parallel vector fields along γ which are perpendicular to γ'. In particular, if Y vanishes at $s=0$, then $E_1=0$. Assuming Y is nontrivial, that is, $E_2\neq 0$, there are no conjugate values. Hence the conjugate locus of a point is empty. Of course, this result is in line with the remark after the proof of Corollary 5.5.4 since we already knew that the cut locus of $\mathbf{R}H^n$ is empty.

$\mathbf{C}P^n$

Owing to Proposition 3.5.1, the geodesics of $\mathbf{C}P^n$ are the projections of the horizontal geodesics of S^{2n+1} with respect to the Riemannian submersion $\pi: S^{2n+1}\to\mathbf{C}P^n$. Let $\tilde\gamma(s)=\cos s\tilde p+\sin s\tilde v$ be a horizontal geodesic of S^{2n+1}, where $\tilde p\in S^{2n+1}$ and $\tilde v\in\mathcal{H}_{\tilde p}$ is a unit vector, and consider the geodesic $\gamma=\pi\circ\tilde\gamma$ of $\mathbf{C}P^n$. It follows that the Jacobi fields along γ are projections of some Jacobi fields along $\tilde\gamma$. Note that whereas a Jacobi field along γ is associated to a variation of $\tilde\gamma$ through horizontal geodesics, this does not imply that the associated Jacobi field along $\tilde\gamma$ must be horizontal. In the following, we want to describe the conjugate points along γ, so we need to describe the Jacobi fields along γ that vanish at $s=0$ and are everywhere orthogonal to γ'.

Consider first the variation through horizontal geodesics

$$\tilde H_0(s,t)=e^{it}\cdot\tilde\gamma(s)=\cos s(\cos t+\sin t(i\tilde p))+\sin s(\cos t+\sin t(i\tilde v)).$$

The associated Jacobi field is

$$\tilde Y_0(s)=i\tilde\gamma(s),$$

and it coincides with the restriction of the vertical vector field (4.5.9) along $\tilde\gamma$. Of course, the corresponding variation of γ is trivial and, accordingly, $\tilde Y_0$ projects down to a trivial Jacobi field along γ.

Next, consider an arbitrary Jacobi field $\tilde Y$ along $\tilde\gamma$ associated to a variation through horizontal geodesics and with the property that it projects down to a Jacobi field Y along γ such that $Y(0)=0$ and $\langle Y,\gamma'\rangle\equiv 0$. We already know that $\tilde Y(s)=\cos s\tilde E_1(s)+\sin s\tilde E_2(s)$ for some parallel vector fields E_1, E_2 along $\tilde\gamma$. The condition that $0=Y(0)=d\pi_{\tilde p}(\tilde Y(0))$ imposes that $\tilde Y(0)$ must be vertical, namely, a multiple of ip. Since $\tilde Y_0$ projects down to zero and the Jacobi fields along a geodesic form a vector space, we can add a suitable multiple of $\tilde Y_0$ to $\tilde Y$ and assume that $\tilde Y(0)=0$. Now $\tilde E_1=0$ and $\tilde Y(s)=\sin s\tilde E_2(s)$. We must have $\langle\tilde Y,\tilde\gamma'\rangle\equiv 0$, so $\tilde E_2(s)$ is a constant vector $\tilde u\in\mathbf{R}^{n+1}$ orthogonal to $\tilde p$ and $\tilde v$. A variation associated to $\tilde Y$ is

$$\tilde H(s,t)=\cos s\,\tilde p+\sin s(\cos t\,\tilde v+\sin t\,\tilde u).$$

Note that $\tilde\gamma_t$ is horizontal if and only if $\tilde\gamma_t'(0)=\cos t\tilde v+\sin t\tilde u$ is orthogonal to $i\tilde p$ if and only if $\tilde u\perp i\tilde p$. We compute

$$\begin{aligned}\langle\tilde Y(s),i\tilde\gamma(s)\rangle &= \langle\sin s\,\tilde u,\cos s(i\tilde p)+\sin s(i\tilde v)\rangle\\ &= \sin^2 s\langle\tilde u,i\tilde v\rangle.\end{aligned}$$

Now there are two cases. If $\tilde{u} \perp i\tilde{v}$, then $\tilde{Y}$ is a horizontal vector field and the corresponding Jacobi field is $Y(s) = \sin s U(s)$, where $U(s)$ is the parallel vector field along γ with $U(0) = d\pi_{\tilde{p}}(\tilde{u})$; the space of such Jacobi fields is $2n-2$-dimensional and the associated conjugate values are multiples of π. On the other hand, if $\tilde{u} = i\tilde{v}$, then the horizontal component of $\tilde{Y}$ is

$$\begin{aligned}\tilde{Y}(s) - \sin^2 s(i\tilde{\gamma}(s)) &= \sin s(i\tilde{v}) - \sin^2 s(\cos s(i\tilde{p}) + \sin s(i\tilde{v})) \\ &= \sin s(\cos s^2(i\tilde{v}) - \sin s\cos s(i\tilde{p})) \\ &= \sin s\cos s(i\tilde{\gamma}'(s)).\end{aligned}$$

In this case, $Y(s) = \sin s\cos s(J\gamma'(s)) = \frac{1}{2}\sin 2s(J\gamma'(s))$; the space of such Jacobi fields is one-dimensional and the associated conjugate values are multiples of $\pi/2$. Finally, it follows from our considerations that the first conjugate locus of a point coincides with the cut locus.

Lie groups

Let G be a Lie group equipped with a bi-invariant metric. In this example, we will describe the conjugate locus of a point in G. By homogeneity, it suffices to compute the conjugate locus of the identity. Denote by $\mathfrak{g}$ the Lie algebra of G. Any geodesic through 1 has the form $\gamma(t) = \exp tX$ for some $X \in \mathfrak{g}$. Let $\{E_1, \ldots, E_n\}$ be a basis of $\mathfrak{g}$. Consider the Jacobi equation $-Y'' + R(\gamma', Y)\gamma' = 0$ along γ. Write $Y(t) = \sum_{i=1}^n y_i(t)E_i$ where y_i are smooth functions on $\mathbf{R}$. Note that $\gamma'(t) = d(L_{\gamma(t)})_1\gamma'(0) = X_{\gamma(t)}$. Then

$$Y'' = \sum_i y_i''E_i + 2y_i'\nabla_X E_i + y_i\nabla_X\nabla_X E_i$$

and

$$R(\gamma', Y)\gamma' = R(X, Y)X = \sum_i y_i\left(\nabla_X\nabla_{E_i}X - \nabla_{[X,E_i]}X\right).$$

A simple calculation using the formula (2.8.8) for the Levi-Cività connection yields that the Jacobi equation along γ has the form

$$\frac{d^2}{dt^2}Y + \mathrm{ad}_X\frac{d}{dt}Y = 0. \tag{5.6.1}$$

Recall that ad_X is a skew-symmetric endomorphism of $\mathfrak{g} \cong T_1G$ with respect to the metric at the identity, so there exists an ad_X-invariant orthogonal decomposition

$$\mathfrak{g} = V_0 \oplus \bigoplus_{j=1}^r V_j$$

where V_0 is the kernel of ad_X and for $j = 1, \ldots, r$ we have $\dim V_j$ is even and the eigenvalues of ad_X on V_j are $\pm i\lambda_j$, $\lambda_j \neq 0$. Now the general solution of (5.6.1) has the form

$$Y(t) = C + Y_0t + \sum_{j=1}^r \cos(\lambda_j t)Y_j + \frac{\sin(\lambda_j t)}{\lambda_j}\mathrm{ad}_X Y_j \tag{5.6.2}$$

where $Y_j \in V_j$ for $j = 0, \ldots, r$ and $C \in \mathfrak{g}$. Therefore the space of Jacobi fields vanishing at $t = 0$ is spanned by

$$Y_0t - Y_j + \cos(\lambda_j t)Y_j + \frac{\sin(\lambda_j t)}{\lambda_j}\mathrm{ad}_X Y_j$$

where $Y_j \in V_j$ for $j = 1, \ldots, r$. This Jacobi field can vanish again only if $Y_0 = 0$; in this case, it is periodic and vanishes exactly when t is a multiple of $2\pi/\lambda_j$. We finally deduce that the points conjugate to 1 along γ are $\gamma(2\pi k/\lambda_j)$, where $k \in \mathbf{Z}$, with multiplicity $\dim V_j$. In particular, the multiplicity of a conjugate point is always even.

5.7 Additional notes

§1 One can recover the results of this chapter by replacing variational calculus by standard calculus on infinite-dimensional smooth manifolds as follows. To begin with, it is necessary to consider a larger class of curves to work with, namely, the absolutely continuous curves $\gamma : [a, b] \to M$ joining p to q with square-integrable $||\gamma'||$. This is a metric space with respect to the distance

$$d(\gamma_1, \gamma_2) = \sup_{t\in[a,b]} d(\gamma_1(t), \gamma_2(t)) + \left(\int_a^b ||\gamma_1'(s) - \gamma_2'(s)||^2 \, ds\right)^{1/2}.$$

Plainly, E and L are continuous functions with respect to this distance. Next, there is a natural way of endowing this space with the structure of a smooth Hilbert manifold. We will not discuss the details of this construction, for which the interested reader is referred to [Kli95, § 2.3] or [PT88, ch. 11]. It turns out that E becomes a smooth function and the first and second variation formulas correspond to its first two derivatives. The main results of this chapter can then be fashioned in the context of Morse theory in Hilbert spaces.

§2 In 1921-30, in the three editions of Blaschke's book [Bla30], it was discussed the problem of whether it is true that a closed surface in $\mathbf{R}^3$ with the property that the first conjugate locus of any point reduces to a single point must be isometric to S^2; he called surfaces with this property *wiedersehens* surfaces. Blaschke studied a number of features of these surfaces and showed, among other things, that: they can be equivalently defined by requiring that the first conjugate point always occurs at the same distance; all of their geodesics are closed and of the same length (hence their name in German); they are homeomorphic to S^2. Of course, if we admit abstract 2-dimensional Riemannian manifolds, then $\mathbf{R}P^2$ also shares this property. In 1963, L. Green [Gre63] proved that S^2 and $\mathbf{R}P^2$ are indeed the only examples. Later, the work of Weinstein [Wei74], Berger-Kazdan [BK80] and Yang [Yan80] extended this result to all dimensions proving that a simply-connected n-dimensional wiedersehens manifold is isometric to S^n.

§3 More generally, it is natural to ask to which extent the conjugate locus structure restricts the topological, differentiable or metric structure of a n-dimensional Riemannian manifold M [War67]. The case of empty conjugate locus will be discussed in the additional notes of chapter 6. The case in which the first tangential conjugate locus of every point $p \in M$ is a round hypersphere in (T_pM, g_p) of the same radius is exactly the subject of §2 above. Consider now the case in which the first tangential conjugate locus of every p is a round sphere in T_pM of the same radius but the multiplicity of the corresponding conjugate points is possibly less than maximal. Namely, we assume that there exists a number $\ell > 0$ and an integer k between 1 and $n-1$ such that, for every $p \in M$ and every geodesic starting at p, the first conjugate point of p occurs at distance ℓ and has multiplicity k; such a manifold is called an *Allamigeon-Warner manifold* [Bes78, chap. 5]. We have already seen that S^n and $\mathbf{C}P^n$ are examples of simply-connected Allamigeon-Warner manifolds; other examples are the quaternionic projective spaces $\mathbf{H}P^n$ and the Cayley projective plane $\mathbf{Ca}P^2$, manifolds that we will discuss later in this book (indeed, we will see that the spheres S^n and the compact projective spaces $\mathbf{R}P^n$, $\mathbf{C}P^n$, $\mathbf{H}P^n$, $\mathbf{Ca}P^2$ are collectively known as the *compact rank one symmetric spaces*). Non-simply-connected examples are given by quotients of those; for instance, $\mathbf{R}P^n$ and lens spaces.

§4 A somehow more specialized condition on a manifold is requiring that the cut-locus structure of each point be similar to that of a compact rank one symmetric space; see [Bes78, chap. 5]. Namely, for distinct points p and q in a complete Riemannian manifold M, the *link from p to q* is the subset $\Lambda(p,q)$ of the unit sphere U_qM of T_qM comprised of vectors of the form $-\gamma'(d(p,q)) \in T_qM$, where $\gamma : [0, d(q,p)] \to M$ is a unit speed minimizing geodesic joining p to q. A compact Riemannian manifold M is called a *Blaschke manifold* if for every $p \in M$ and $q \in \mathrm{Cut}(p)$, the link $\Lambda(p,q)$ is a great sphere of U_qM; here it is not required that the tangential cut-locus at a point is a round sphere, but this follows from the definition. It is known that a Blaschke manifold is Allamigeon-Warner, and both concepts are equivalent in the simply-connected case. Note that $\Lambda(p,q)$ equals U_qM for S^n, it consists of two antipodal points of U_qM for $\mathbf{R}P^n$, and it consists of a great circle of U_qM for $\mathbf{C}P^n$. One sees that $\Lambda(p,q)$ is a great 3-sphere of U_qM for $\mathbf{H}P^n$ and a great 7-sphere of U_qM for $\mathbf{Ca}P^2$. The *Blaschke conjecture* asserts that every Blaschke manifold is isometric to a compact rank one symmetric space. This is one of the famous yet open problems in geometry, with many partial results proved. The book [Bes78] contains a discussion of this conjecture as well as more general discussions of Riemannian manifolds all of whose geodesics are closed; see [Rez94] for a more recent bibliography.

5.8 Exercises

1 Let $\gamma : [a,b] \to M$ be a geodesic parametrized with unit speed in a Riemannian manifold M, and let H be a piecewise smooth variation of γ with associated variational vector field Y. Show that

$$\begin{aligned}\frac{d^2}{dt^2}\Big|_{t=0} L(\gamma_t) &= \langle \overline{\nabla}_{\frac{\partial}{\partial t}} \frac{\partial}{\partial t}\Big|_{t=0}, \gamma' \rangle \Big|_a^b + \int_a^b ||Y'||^2 + \langle R(\gamma', Y)\gamma', Y\rangle - \langle Y', \gamma'\rangle^2 \, ds \\ &= \langle \overline{\nabla}_{\frac{\partial}{\partial t}} \frac{\partial}{\partial t}\Big|_{t=0}, \gamma' \rangle \Big|_a^b + \int_a^b ||Y_\perp'||^2 + \langle R(\gamma', Y_\perp)\gamma', Y_\perp\rangle \, ds,\end{aligned}$$

where $Y_\perp = Y - \langle Y, \gamma'\rangle\gamma'$ is the normal component of Y.

2 Let $\gamma : [0,\ell] \to M$ be a geodesic in a Riemannian manifold M. Consider the index form I on the space of piecewise smooth vector fields along γ vanishing at 0 and ℓ. Prove that the kernel of I consists precisely of the Jacobi fields along γ vanishing at 0 and ℓ. (Hint: Use the formula (5.4.1), and for a given element Y in the kernel of I, choose suitable elements X as it was done in the proof of Proposition 5.3.5).

3 Let $\gamma : [0,\ell] \to M$ be a geodesic in a Riemannian manifold M. Extend the definition of the index form I to the space of piecewise smooth vector fields along γ non-necessarily vanishing at the endpoints. Prove that if γ is a minimizing geodesic, X is a smooth vector field along γ, and Y is a Jacobi vector field along γ with the same values as X at the endpoints, then $I(X,X) \geq I(Y,Y)$.

4 Let N_1 and N_2 be two closed submanifolds of a complete Riemannian manifold M. Assume that one of N_1, N_2 is compact.

a. Prove that there exist points $p_1 \in N_1$ and $p_2 \in N_2$ such that $d(N_1,N_2) = d(p_1,p_2)$.
b. Prove that there exists a geodesic γ of M joining p_1 and p_2 and that $L(\gamma) = d(N_1,N_2)$.
c. Prove that γ is perpendicular to N_1 (resp. N_2) at p_1 (resp. p_2). (Hint: Use the first variation formula.)

5 Let $\gamma : [a,b] \to M$ be a geodesic in a Riemannian manifold, and let $\gamma(a) = p$ and $\gamma(b) = q$. Prove that if p and q are not conjugate along γ, then given $u \in T_pM$ and $v \in T_qM$, there exists a unique Jacobi field J along γ such that $J(a) = u$ and $J(b) = v$.

6 Let M be a Riemannian manifold, and let X be a Killing field on M.

a. If γ is a geodesic in M, prove that the restriction $J = X \circ \gamma$ of X to a vector field along γ is a Jacobi field.

b. If M is complete and $p \in M$, prove that X is completely determined by the values of $X(p) \in T_pM$ and $(\nabla X)_p \in \mathrm{End}(T_pM)$.

c. Deduce from part (b) that the dimension of the Lie algebra of Killing fields on M is bounded by $\frac{1}{2}n(n+1)$, where $n = \dim M$.

7 Let M be a Riemannian manifold and let X be a Killing field on M. Prove that

$$\nabla_U \nabla_V X - \nabla_{\nabla_U V} X + R(X,U)V = 0$$

for all smooth vector fields U and V on M. (Hint: Use Exercise 6(a).)

8 Let (M,g) be a Riemannian manifold, fix $p \in M$ and choose an orthonormal basis $\{e_1, \ldots, e_n\}$ of T_pM. Let $\epsilon > 0$ be such that $\exp_p : B(0_p, \epsilon) \subset T_pM \to M$ is a diffeomorphism onto its image U, and use it to define a local coordinates $x^1, \ldots, x^n$ around p. Let $v \in T_pM$ be a unit vector and consider the geodesic $t \mapsto \exp_p(tv)$. Show that the coefficients of the metric in this chart admit expansions

$$g_{ij}(\exp_p tv) = \delta_{ij} + \langle R(v,e_i)v, e_j \rangle \frac{t^2}{3} + \mathrm{O}(t^3),$$

where $1 \leq i,j \leq n$, $0 < t < \epsilon$, and $\mathrm{O}(t^3)$ denotes a term such that $\mathrm{O}(t^3)/t^2 \to 0$ as $t \to 0$. (Hint: Use the result of Scholium 5.4.5.)

9 Let (M,g) be a compact Riemannian manifold.

a. Prove that if the Ricci tensor of M is negative definite everywhere, then the isometry group $\mathrm{Iso}(M,g)$ is finite. (Hint: Use exercise 7 and the divergence theorem (exercise 14 in chapter 4) to show that there are no nontrivial Killing fields on M.)

b. Prove that if the Ricci tensor of M is negative semi-definite everywhere, then any Killing field is parallel.

10 Let G be a Lie group equipped with a bi-invariant metric. Use exercise 12 of chapter 2 and exercise 6(a) above to show that the restriction of a left-invariant or right-invariant vector field along a geodesic γ is a Jacobi field. Check that not every Jacobi field along γ has the form $J_1 + J_2$, where $J_1 = X_1 \circ \gamma$, $J_2 = X_2 \circ \gamma$, X_1 is left-invariant and X_2 is right-invariant.

11 Let M be a Riemannian manifold.

a. Prove that the "cut-distance" function $\rho : UM \to (0, +\infty]$ is upper semi-continuous. (Hint: For $v_i \to v$, prove that $\limsup \rho(v_i) \leq \rho(v)$ using the continuity of the distance function d on $M \times M$).

b. Assume now that M is complete and prove that ρ is continuous. (Hint: for $v_i \to v$, prove that $\liminf \rho(v_i) \geq \rho(v)$ using ideas from the proof of Proposition 5.5.5.)

c. Deduce from part (b) that the injectivity radius inj_p depends continously on p.

Applications

6.1 Introduction

In this chapter, we collect a few basic and important theorems of Riemannian geometry that we prove by using the concepts introduced so far. We also introduce some other important techniques along the way.

We start by discussing manifolds of constant curvature. If one agrees that curvature is the main invariant of Riemannian geometry, then in some sense the spaces of constant curvature should be the simplest models of Riemannian manifolds. It is therefore very natural to try to understand those manifolds. Since curvature is a local invariant, one can only expect to get global results by further imposing other topological conditions.

Next we turn to the relation between curvature and topology. This a central and recurring theme for research in Riemannian geometry. One of its early pioneers was Heinz Hopf in the 1920's who asked to what extent the existence of a Riemannian metric with particular curvature properties restricts the topology of the underlying smooth manifold. Since then the subject has expanded so much that the scope of this book can only afford a glimpse at it.

It is worthwhile pointing out that not only the theorems in this chapter are part of a central core of results in Riemannian geometry, but also the arguments and techniques in the proofs can be applied in more general contexts to a wealth of other important problems in geometry.

6.2 Space forms

A complete Riemannian manifold with constant curvature is called a *space form*. If M is a space form, its universal Riemannian covering manifold $\tilde{M}$ is a simply-connected space form by Proposition 3.3.8. Moreover, M is isometric to $\tilde{M}/\Gamma$ with the quotient metric, where Γ is a free and proper discontinuous subgroup of isometries of $\tilde{M}$, see section 1.3. So the classification of space forms can be accomplished in two steps, as follows:

a. Classification of the simply-connected space forms.

b. For each simply-connected space form, classification of the subgroups of isometries acting freely and properly discontinuously.

In this section, we will prove the Killing-Hopf theorem that solves part (a) in this program. Despite a lot being known about part (b), it is yet an unsolved problem, and we include a brief discussion about it after the proof of the theorem.

We first prove a local result.

6.2.1 Theorem *Fix $k \in \mathbf{R}$. Then any two Riemannian manifolds of constant curvature k of the same dimension are locally isometric.*

Proof. Let M, $\tilde{M}$ be two Riemannian manifolds of constant curvature k. Fix points $p \in M$, $\tilde{p} \in \tilde{M}$ and choose a linear isometry $f : T_pM \to T_{\tilde{p}}\tilde{M}$. Choose open balls $U \subset T_pM$, $\tilde{U} \subset T_{\tilde{p}}\tilde{M}$ with $\tilde{U} = f(U)$ that determine normal neighborhoods $V = \exp_p(U)$, $\tilde{V} = \exp_{\tilde{p}}(\tilde{U})$. Now we have a diffeomorphism $F : V \to \tilde{V}$ given by

$$\begin{array}{ccc} U & \xrightarrow{f} & \tilde{U} \\ {\scriptstyle \exp_p}\downarrow & & \downarrow{\scriptstyle \exp_{\tilde{p}}} \\ V & \xrightarrow[F]{} & \tilde{V} \end{array}$$

namely, $F \circ \exp_p = \exp_{\tilde{p}} \circ f$. Note that $F(p) = \tilde{p}$ and $dF_p = f$. We shall prove that F is an isometry.

We need to prove that $dF_q : T_qM \to T_{\tilde{q}}\tilde{M}$ is a linear isometry, where $q \in V$ is arbitrary and $\tilde{q} = F(q)$. Write $q = \gamma_v(t_0)$ where γ_v is the radial geodesic from p with initial unit velocity $v \in T_pM$ and $t_0 \in [0, \epsilon)$. We orthogonally decompose $T_qM = \mathbf{R}\gamma_v'(t_0) \oplus W$, where W is the orthogonal complement, and similarly $T_{\tilde{q}}\tilde{M} = \mathbf{R}\gamma_{\tilde{v}}'(t_0) \oplus \tilde{W}$, where $\tilde{v} = f(v)$.

Note $F \circ \gamma_v$ is the geodesic $\gamma_{\tilde{v}}$ in $\tilde{M}$, so by the chain rule

$$||dF_q(\gamma_v'(t_0))|| = ||\gamma_{\tilde{v}}'(t_0)|| = ||\tilde{v}|| = ||v|| = ||\gamma_v'(t_0)||.$$

Furthermore, by the Gauss lemma 5.5.1 (or 3.2.1), $d(\exp_p)_{t_0v} : T_pM \to T_qM$ sends the orthogonal decomposition $T_pM = \mathbf{R}v \oplus (\mathbf{R}v)^\perp$ to the orthogonal decomposition $T_qM = \mathbf{R}\gamma_v'(t_0) \oplus W$, and similarly for $d(\exp_{\tilde{p}})_{t_0\tilde{v}}$. It follows that dF_q sends the orthogonal decomposition $T_qM = \mathbf{R}\gamma_v'(t_0) \oplus W$ to $T_{\tilde{q}}\tilde{M} = \mathbf{R}\gamma_{\tilde{v}}'(t_0) \oplus \tilde{W}$. It remains only to check that dF_q restricts to an isometry $W \to \tilde{W}$.

It is here and only here that we use the assumption on the sectional curvatures. Let $u \in T_pM$ be orthogonal to v and let $\tilde{u} = f(u) \in T_{\tilde{p}}\tilde{M}$. Extend u, $\tilde{u}$ to parallel vector fields U, $\tilde{U}$ along γ_v, $\gamma_{\tilde{v}}$, respectively. On one hand, the Jacobi fields Y, $\tilde{Y}$ along γ_v, $\gamma_{\tilde{v}}$, resp., with initial conditions $Y(0) = \tilde{Y}(0) = 0$, $Y'(0) = u$, $\tilde{Y}'(0) = \tilde{u}$ are given by $Y(t) = d(\exp_p)_{tv}(tu)$, $\tilde{Y}(t) = d(\exp_{\tilde{p}})_{t\tilde{v}}(t\tilde{u})$, due to Scholium 5.4.5. On the other hand, the Jacobi equation along a geodesic in a space of constant curvature k is given by $Y'' + kY = 0$. It follows that

$$Y(t) = \begin{cases} \frac{\sin(\sqrt{k}t)}{\sqrt{k}}U(t), & \text{if } k > 0, \\ \frac{\sinh(\sqrt{-k}t)}{\sqrt{-k}}U(t), & \text{if } k < 0, \end{cases} \quad \text{and} \quad \tilde{Y}(t) = \begin{cases} \frac{\sin(\sqrt{k}t)}{\sqrt{k}}U(t), & \text{if } k > 0, \\ \frac{\sinh(\sqrt{-k}t)}{\sqrt{-k}}U(t), & \text{if } k < 0, \end{cases}$$

and

$$Y(t) = tU(t) \quad \text{and} \quad \tilde{Y}(t) = t\tilde{U}(t)$$

if $k = 0$. In any case

$$||\tilde{Y}(t)|| = ||Y(t)||.$$

Since $Y(t_0) \in W$ is an arbitrary vector and

$$\begin{aligned} dF_q(Y(t_0)) &= dF_q\left(d(\exp_p)_{t_0v}(t_0u)\right) \\ &= d(\exp_{\tilde{p}})_{t\tilde{v}}(t_0f(u)) \\ &= \tilde{Y}(t_0), \end{aligned}$$

it follows that $dF_q : W \to \tilde{W}$ is an isometry, and this finishes the proof. □

If (M, g) is a space form of curvature k, then, for a positive real number λ, $(M, \lambda g)$ is a space form of curvature $\lambda^{-1}k$, see Exercise 2 in chapter 4. Therefore, the metric g can be normalized so that k becomes equal to one of 0, 1, or -1.

6.2.2 Theorem (Killing-Hopf) *Let M be a simply-connected space form of curvature k and dimension bigger than one. Then M is isometric to:*

a. the Euclidean space $\mathbf{R}^n$, if $k = 0$;
b. the real hyperbolic space $\mathbf{R}H^n$, if $k = -1$;
c. the unit sphere S^n, if $k = 1$.

Proof. Let $\tilde{M}$ be $\mathbf{R}^n$, $\mathbf{R}H^n$ or S^n according to whether $k = 0$, -1 or 1. Fix $\tilde{p} \in \tilde{M}$, $p \in M$ and choose a linear isometry $f : T_{\tilde{p}}\tilde{M} \to T_pM$. As in the proof of Theorem 6.2.1, this data can be used to define an isometry $F : \tilde{V} \to V$ with $F(\tilde{p}) = p$ and $dF_{\tilde{p}} = f$, where V, $\tilde{V}$ are certain normal neighborhoods of p, $\tilde{p}$. We shall see that F can be extended to an isometry $\tilde{M} \to M$.

Consider first the case $k = 0$ or -1. Since the cut locus of a point in $\mathbf{R}^n$ or $\mathbf{R}H^n$ is empty, we can take $\tilde{V} = \tilde{M}$ as a normal neighborhood, and using the completeness of M, extend F to a map $\tilde{M} \to M$ by the same formula, namely, $F \circ \exp_{\tilde{p}} = \exp_p \circ f$. Note, however, that in principle F does not have to be a diffeomorphism, because $f(T_{\tilde{p}}\tilde{M}) = T_pM$ does not in principle exponentiate to a normal neighborhood of p. Nevertheless, the proof of Theorem 6.2.1 (using the global Gauss lemma 5.5.1) carries through to show that F is a local isometry. Since $\tilde{M}$ is complete, Proposition 3.3.8(b) can be applied to yield that F is a Riemannian covering map and hence, since M is assumed to be simply-connected, F must be an isometry.

Consider now $k = 1$. Here the above argument yields a local isometry $F : \tilde{V}_{\tilde{p}} \to M$, where $\tilde{V}_{\tilde{p}} = S^n \setminus \{-\tilde{p}\}$ is the maximal normal neighborhood of $\tilde{p}$. To finish, we choose another point $\tilde{q} \in S^n \setminus \{\tilde{p}, -\tilde{p}\}$ and construct a similar local isometry $G : \tilde{V}_{\tilde{q}} \to S^n$, with initial data $G(\tilde{q}) = F(\tilde{q})$ and $dG_{\tilde{q}} = dF_{\tilde{q}}$, where $V_{\tilde{q}} = S^n \setminus \{-\tilde{q}\}$. By exercise 15 of chapter 3, F and G can be pasted together to define a local isometry $S^n \to M$. The rest of the proof is as above, using the completeness of S^n and the simple-connectedness of M. □

Depending on the context in which one is interested, it is possible to find in the literature other proofs of Theorem 6.2.2 different from the above one. The argument that we chose to use, based on Jacobi fields, works in a more general context, and will be used to prove a generalization of this theorem in chapter ??? of part 2.

Next, we discuss the case of non-simply-connected space forms. In the flat case, the main result is the following theorem.

6.2.3 Theorem (Bieberbach) *A compact flat manifold M is finitely covered by a torus.*

Namely, Bieberbach showed that the fundamental group $\pi_1(M)$ contains a free Abelian normal subgroup Γ of rank $n = \dim M$ and finite index, so there is a finite covering

$$\pi_1(M)/\Gamma \to \mathbf{R}^n/\Gamma \to \mathbf{R}^n/\pi_1(M) = M.$$

(For an example, review the contents of exercise 10 of chapter 1.) The complete classification of compact flat Riemannian manifolds is known only in the cases $n = 2$, 3; see [Wol84, Cha86] for proofs of Bieberbach's theorem and these classifications.

Next we consider non-simply-connected space forms of positive curvature. In even dimensions, the only examples are the real projective spaces, as the following result shows.

6.2.4 Theorem *An even-dimensional space form of positive curvature is isometric either to S^{2n} or to $\mathbf{R}P^{2n}$.*

Proof. We know that $M = S^{2n}/\Gamma$, where Γ is a subgroup of $\mathbf{O}(2n+1)$ acting freely and properly discontinuously on S^{2n}. Since this action is free, if an element of Γ admits a $+1$-eigenvalue then it must be the identity id. Recall that the eigenvalues of an orthogonal transformation are unimodular complex numbers, and the non-real ones must occur in complex conjugate pairs.

Next, let $\gamma \in \Gamma$. Then $\gamma^2 \in \mathbf{SO}(2n+1)$, and since $2n+1$ is odd, γ^2 admits an eigenvalue $+1$, thus $\gamma^2 = \mathrm{id}$. This implies that all the eigenvalues of γ are ± 1. If $\gamma \neq \mathrm{id}$, it follows that all the eigenvalues of γ are -1, namely, $\gamma = -\mathrm{id}$. Hence $\Gamma = \{\mathrm{id}\}$ or $\Gamma = \{\pm\mathrm{id}\}$. □

The odd-dimensional space forms of positive curvature have been completely classified by J. Wolf [Wol84]. Here we just present a very rich family of examples.

6.2.5 Example (Lens spaces) Let p, q be relatively prime integers. The *lens space* $L_{p;q}$ is the quotient Riemannian manifold S^3/Γ, where we view

$$S^3 = \{ (z_1, z_2) \in \mathbf{C}^2 \mid |z_1|^2 + |z_2|^2 = 1 \},$$

and Γ is the cyclic group of order p generated by the element

$$t_{p;q}(z_1, z_2) = (\omega z_1, \omega^q z_2),$$

where ω is a pth root of unity. Note that $L_{2;1} = \mathbf{R}P^3$. More generally, let $q_2, \ldots, q_n$ be integers relatively prime to an integer p. The *lens space* $L_{p;q_2,\ldots,q_n}$ is the quotient Riemannian manifold S^{2n-1}/Γ, where we view

$$S^{2n-1} = \{ (z_1, \ldots, z_n) \in \mathbf{C}^2 \mid |z_1|^2 + \cdots + |z_n|^2 = 1 \},$$

and Γ is the cyclic group of order p generated by the element

$$t_{p;q_2,\ldots,q_n}(z_1, z_2, \ldots, z_n) = (\omega z_1, \omega^{q_2} z_2, \ldots, \omega^{q_n} z_n).$$

Of course, a lens space is a non-simply-connected space form of positive curvature. The 3-dimensional lens spaces were introduced by Tietze in 1908. In general, lens spaces are important in topology because they provide examples of non-homeomorphic compact manifolds which are homotopy-equivalent (see [Mun84, §40, §69]). Historywise they can thus be seen as representing the birth of geometric topology of manifolds as distinct from algebraic topology. ⋆

A space form of negative curvature is called a *hyperbolic manifold*. Of course, a hyperbolic manifold is isometric to the quotient of $\mathbf{R}H^n$ by a group of isometries Γ acting freely and proper discontinously. A compact orientable surface of genus $g \geq 2$ admits many hyperbolic metrics, which are constructed as follows. It is a theorem of Radó [Rad24] that any compact surface is homeomorphic to the identification space of a polygon whose sides are identified in pairs. In particular, a compact orientable surface S_g of genus g is realized as a regular $4g$-sided polygon P with a certain identification of the sides. The vertices of P are all identified to one point, so in order to get a smooth surface it is necessary that the sum of the inner angles of P be 2π. Note that P cannot be taken to be an Euclidean polygon, for in that case the sum of the inner angles is known to be $(4g-2)\pi > 2\pi$ for $g \geq 2$. Instead, we construct P as a regular polygon in the ball model $\mathbf{B}^2$ of $\mathbf{R}H^2$ having the center at $(0,0)$ and with the sides being geodesic segments. In this case, by the Gauss-Bonnet theorem the sum of the inner angles is $(4g-2)\pi - A$, where A denotes the area of P. It is clear that there exist such polygons in $\mathbf{D}^2$ with arbitrary diameter, and that A varies continuously with the diameter, between zero (when the diameter is near zero) and $(4g-2)\pi$ (when the angles are near zero). Since $(4g-2)\pi > 2\pi$, it follows from the intermediate value theorem that it is possible to construct P such that the sum of the inner angles is 2π. Next one sees that the identifications between pairs of sides can be realized by isometries of $\mathbf{D}^2$ such that these isometries generate a discrete subgroup Γ of the isometry group of $\mathbf{D}^2$ acting freely and properly discontinuously. This shows that $S_g = \mathbf{D}^2/\Gamma$ admits a hyperbolic metric. Further, it is known that the hyperbolic metric on S_g for $g \geq 2$ is not unique. It is a classical result that there exist natural bijections between the following sets of structures on a compact oriented surface S_g: conformal classes of Riemannian metrics; complex structures compatible with the orientation; hyperbolic metrics (see e.g. [Jos06]). The *moduli space* $\mathcal{M}_g$ of S_g is the space of equivalence classes of hyperbolic metrics on S_g, where two hyperbolic metrics belong to the same class if and only if they differ by a diffeomorphism of S_g. It turns out that $\mathcal{M}_g$ is not a manifold: singularities develop exactly at the hyperbolic metrics admitting nontrivial isometry groups. For this reason, Teichmüller introduced a weaker equivalence relation on the space of hyperbolic metrics on S_g by requiring two of them to be equivalent if they differ by a diffeomorphism which is homotopic to the identity; the *Teichmüller space* $\mathcal{T}_g$ of S_g is the resulting space of equivalence classes. It is known that $\mathcal{T}_g$ admits the structure of a smooth manifold of dimension $6g-6$ if $g \geq 2$ [EE69].

In the higher dimensional case, it is much more difficult to construct hyperbolic metrics, and most of the progress in this direction has been made in the 3-dimensional case, see [Thu97].

6.3 Synge's theorem

We will use the following lemma in the proofs of Synge's and Preissmann's theorems. It is easy to see that the compactness assumption in it is essential.

6.3.1 Lemma (Cartan) *Let M be a compact Riemannian manifold. Assume that M is not simply-connected. Then every nontrivial free homotopy class $\mathcal{C}$ of loops contains a closed geodesic of minimal length in $\mathcal{C}$.*

Proof. We first claim that since M is compact, it is possible to find $\epsilon > 0$ such that any two points of M within distance less than ϵ can be joined by a unique minimizing geodesic, and this geodesic depends smoothly on its endpoints. Indeed, cover M by finitely many balls $B(p_i, \epsilon_i/2)$ where $p_i \in M$, $\epsilon_i > 0$, and $B(p_i, \epsilon_i)$ is a δ_i-totally normal ball for some $\delta_i > 0$ as in Proposition 2.4.7, for $i = 1, \ldots, k$. Take $\epsilon = \min_i\{\frac{1}{2}\epsilon_i, \delta_i\}$. If $d(x, y) < \epsilon$ for points x, $y \in M$, then $x \in B(p_{i_0}, \epsilon_{i_0}/2)$ for some i_0, and then

$$d(y, p_{i_0}) \leq d(y, x) + d(x, p_{i_0}) < \epsilon + \frac{\epsilon_{i_0}}{2} \leq \epsilon_{i_0}.$$

Hence x, $y \in B(p_{i_0}, \epsilon_{i_0})$ with $d(x, y) < \delta_{i_0}$, so the claim follows from the quoted proposition.

Let ℓ be the infimum of the lengths of the piecewise smooth curves in $\mathcal{C}$, and take a minimizing sequence (η_j) in $\mathcal{C}$ such that each η_j is parametrized on $[0, 1]$ with constant speed. Since (η_j) is a minimizing sequence, $L = \sup_j L(\eta_j)$ is finite. Choose a subdivision $0 = t_0 < t_1 < \ldots < t_n = 1$ with $t_i - t_{i-1} < \epsilon/2L$ for $i = 1, \ldots, n$. Then

$$d(\eta_j(t_{i-1}), \eta_j(t)) \leq \int_{t_{i-1}}^{t} ||\eta_j'(t)||\, dt \leq L(t_i - t_{i-1}) < \frac{\epsilon}{2}$$

for $t_{i-1} \leq t \leq t_i$. This estimate allows us to replace each curve η_j by the broken geodesic γ_j joining the points $\eta_j(0), \eta_j(t_1), \ldots, \eta_j(1)$. For every j, γ_j is homotopic to η_j; this can be seen as follows. Owing to

$$d(\gamma_j(t), \eta_j(t)) \leq d(\gamma_j(t), \gamma_j(t_{i-1})) + d(\eta_j(t_{i-1}), \eta_j(t)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for $t_{i-1} \leq t \leq t_i$, we can construct a smooth homotopy from $\eta_j|_{[t_{i-1}, t_i]}$ into $\gamma_j|_{[t_{i-1}, t_i]}$ by using the shortest geodesic from $\eta_j(t)$ to $\gamma_j(t)$.

It is clear that $L(\gamma_j) \leq L(\eta_j)$, so (γ_j) is also a minimizing sequence in $\mathcal{C}$. Using again the compactness of M, we can select a subsequence of (γ_j), denoted by the same symbol, such that $(\gamma_j(t_i))$ converges to a point p_i as $j \to \infty$ for all i. It follows that (γ_j) converges in the C^1-topology to the broken geodesic γ joining the p_i. It is clear that γ belongs to $\mathcal{C}$ and has length ℓ. Since γ is of minimal length in $\mathcal{C}$, it is locally minimizing. By Theorem 3.2.6, γ is a geodesic. □

In the case of a simply connected compact Riemannian manifold, it is still true that there exists at least one closed geodesic (Lyusternik-Fet [LF51]). More specifically, in the case of S^2, it is known that every Riemannian metric must admit at least 3 geometrically distinct simple closed geodesics (Lyusternik-Schnirelmann [LŠ47]).

6.3.2 Theorem (Synge) *An even-dimensional orientable compact Riemannian manifold M of positive sectional curvature must be simply connected.*

We remark that each one of the hypotheses in the statement of Synge's theorem is essential. In fact, the following manifolds are not simply-connected: $\mathbf{R}P^2$ is even-dimensional, compact and positively curved, but nonorientable; $\mathbf{R}P^3$ is compact, orientable and positively curved, but odd-dimensional; a flat 2-torus is even-dimensional, compact and orientable, but not positively curved; a 2-sphere with two punctures is even-dimensional, orientable and positively curved, but non-compact.

Proof of Theorem 6.3.2. Suppose, on the contrary, that M is not simply-connected and let $\mathcal{C}$ denote a nontrivial free homotopy class of loops. By Lemma 6.3.1, there exists a closed geodesic $\gamma : [0, \ell] \to M$, parametrized with unit speed, such that $L(\gamma) = \ell = \inf_{\eta \in \mathcal{C}} L(\eta)$. Let $p = \gamma(0) = \gamma(\ell)$, and denote by $P : T_pM \to T_pM$ the parallel translation map along γ from 0 to ℓ. Fix an orientation of M. Since the parallel translation maps along γ from 0 to t, for $0 \leq t \leq \ell$, join P to the identity map of T_pM, we have that P is orientation-preserving. Since γ is a geodesic, $\gamma'(0)$ is a fixed vector of P. Now P, being an isometry, leaves the orthogonal complement $\langle \gamma'(0) \rangle^\perp$ invariant. Since the dimension of this subspace is odd, it contains a nonzero vector y that is fixed under P. Let Y be the parallel vector field along γ that extends y, and construct a variation $\{\gamma_t\}$ of γ through closed curves with associated variational vector field given by Y. Since M is positively curved, $\langle R(Y, \gamma')Y, \gamma' \rangle < 0$. Using the variation formulas (5.3.3) and (5.3.9), we get that

$$\left.\frac{d}{dt}\right|_{t=0} E(\gamma_t) = 0 \quad \text{and} \quad \left.\frac{d^2}{dt^2}\right|_{t=0} E(\gamma_t) < 0.$$

Then, for t sufficiently small, we have that $E(\gamma_t) < E(\gamma)$ and

$$L(\gamma_t)^2 \leq 2\ell\ E(\gamma_t) < 2\ell\, E(\gamma) = L(\gamma)^2,$$

and this contradicts the fact that γ is of minimal length in $\mathcal{C}$. Hence $\mathcal{C}$ cannot exist and M is simply-connected. □

6.3.3 Corollary *An even-dimensional compact Riemannian manifold M of positive sectional curvature has fundamental group of order at most two.*

Proof. We may assume M is non-orientable. Let $\tilde{M}$ be the orientable double cover of M. Then $\tilde{M}$ is connected and satisfies the hypotheses of Synge's theorem 6.3.2, so it is simply connected. The result follows. □

It follows from Corollary 6.3.3 that there exists no Riemannian metric of positive sectional curvature in $\mathbf{R}P^m \times \mathbf{R}P^n$ if $m + n$ is even. Indeed, otherwise this manifold would satisfy the hypotheses of the corollary but its fundamental group is isomorphic to $\mathbf{Z}_2 \oplus \mathbf{Z}_2$. It is interesting to compare this example with the fact that the nonexistence of a positively curved Riemannian metric in $S^2 \times S^2$ is still an unsettled question (see Add. note 4).

6.4 Bonnet-Myers' theorem

The following result is an elementary example of a comparison theorem in Riemannian geometry. Note that the right-hand side in (6.4.2) is exactly the Ricci curvature of the sphere $S^n(R)$.

6.4.1 Theorem (Bonnet-Myers) *Let M be a complete Riemannian manifold of dimension n. Assume there exists a constant $R > 0$ such that*

$$\text{Ric}(v, v) \geq \frac{n-1}{R^2}\, g(v, v) \tag{6.4.2}$$

for every $v \in TM$. Then

$$\text{diam}(M) \leq \text{diam}(S^n(R)) = \pi R.$$

In particular, M is compact and has finite fundamental group $\pi_1(M)$.

Proof. Recall that $\text{diam}(M) = \sup\{\, d(x, y) \mid x,\ y \in M\,\}$. We will show that the distance of two given points p, $q \in M$ is bounded above by πR. Since M is complete, there exists a minimal geodesic $\gamma : [0, L] \to M$ with unit speed and such that $\gamma(0) = p$ and $\gamma(L) = q$. Because γ is minimal, $I(Y, Y) \geq 0$ for all vector fields Y along γ vanishing at the endpoints. We will use this remark below for some suitable vector fields.

Select an orthonormal basis $\{e_1, \ldots, e_n\}$ of T_pM with $e_1 = \gamma'(0)$, and extend it to parallel orthonormal frame $\{E_1, \ldots, E_n\}$ along γ; of course, $E_1 = \gamma'$. Set

$$Y_i(s) = \sin\frac{\pi s}{L} E_i(s)$$

for $i = 2, \ldots, n$. Then

$$\begin{aligned} I(Y_i, Y_i) &= \int_0^L -\langle Y_i'', Y_i\rangle + \langle R(\gamma', Y_i)\gamma', Y_i\rangle\, ds \\ &= \int_0^L \sin^2\frac{\pi s}{L}\left(\frac{\pi^2}{L^2} + \langle R(\gamma', E_i)\gamma', E_i\rangle\right) ds. \end{aligned}$$

Noting that each Y_i vanishes at the endpoints of γ, we have

$$\begin{aligned} 0 \leq \sum_{i=2}^{n} I(Y_i, Y_i) &= \int_0^L \sin^2\frac{\pi s}{L}\left((n-1)\frac{\pi^2}{L^2} - \text{Ric}(\gamma', \gamma')\right) ds \\ &\leq (n-1)\left(\frac{\pi^2}{L^2} - \frac{1}{R^2}\right)\int_0^L \sin^2\frac{\pi s}{L}\, ds, \end{aligned}$$

using the assumption on the Ricci curvature. This proves that $d(p, q) = L \leq \pi R$. We conclude that $\text{diam}(M) \leq \pi R$.

The other assertions in the statement can now be easily verified. The manifold M is complete and bounded, thus, in view of Corollary 3.3.7, compact. Let $\tilde{M}$ denote the Riemannian universal covering manifold of M. Since $\tilde{M}$ is complete and satisfies the same estimate on the Ricci curvature as M, the previous results imply that $\tilde{M}$ is compact, forcing $\pi_1(M)$ to be finite. This completes the proof of the theorem. □

6.4.3 Corollary *No compact nontrivial product manifold $S^1 \times M$ admits a metric of positive Ricci curvature.*

6.4.4 Remark The assumption about the Ricci curvature in the statement of the Bonnet-Myers theorem cannot be relaxed in the sense of requiring that the Ricci curvature only be positive, as the following example shows. The two-sheeted hyperboloid

$$\{ (x, y, z) \in \mathbf{R}^3 \mid x^2 + y^2 - z^2 = -1 \}$$

with the metric induced from $\mathbf{R}^3$ is complete, non-compact, and has Gaussian curvature at a point (x, y, z) given by $(x^2 + y^2 + z^2)^{-2}$, which, despite being positive, goes to zero as the point tends to infinity. ⋆

6.5 Nonpositively curved manifols

One of the main features of nonpositively curved manifols is the abundance of convex functions. Recall that a continuous function $f : I \to \mathbf{R}$ defined on an interval I is called convex if $f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$ for every $t \in (0, 1)$ and x, $y \in I$. If f is of smooth, this condition is equivalent to requiring that its second derivative $f'' \geq 0$. In the case of a continous function f on a complete Riemannian manifold M, we say that f is *convex* if its restriction $f \circ \gamma$ is convex for every geodesic γ of M. Strict convexity is defined analogously by replacing the inequalities above by the strict inequalities. Our point of view in this section is that most of the important results about the geometry of manifolds with nonpositive curvature can be derived by using appropriate convex functions on the manifold.

We will use the following remark in the proof of Lemma 6.5.1. If a convex function admits two global minima, then a geodesic connecting these two points also consists of global minima of the function. In fact, the function restricted to the geodesic is convex, and this implies that it cannot have bigger values on the interior of the segment than at the endpoints forcing it to be constant along the geodesic segment. A similar argument shows that any local minimum of a convex function on a complete Riemannian manifold must in fact be a global one.

6.5.1 Lemma *Let γ be a geodesic in a Riemannian manifold M. If the sectional curvature along γ is nonpositive, then there are no conjugate points along γ.*

Proof. Let Y be a Jacobi field along γ. We claim that the fuction $f = ||Y||^2$ is convex. In order to prove this, we recall the Jacobi equation $-Y'' + R(\gamma', Y)\gamma' = 0$ and differentiate f twice to get

$$\begin{aligned} f'' &= 2(\langle Y'', Y\rangle + ||Y'||^2) \\ &= 2(\langle R(\gamma', Y)\gamma', Y\rangle + ||Y'||^2) \\ &\geq 0, \end{aligned}$$

in view of the assumption on the curvature; this proves the claim. Now if $f(t_1) = f(t_2) = 0$ for some $t_1 < t_2$, then $f|_{[t_1,t_2]} \equiv 0$, whence Y is trivial. Hence there are no conjugate points along γ. □

6.5.2 Theorem (Hadamard-Cartan) *Let M be a complete Riemannian manifold with nonpositive sectional curvature. Then, for every point $p \in M$, the exponential map $\exp_p : T_pM \to M$ is a smooth covering. In particular, M is diffeomorphic to $\mathbf{R}^n$ if it is simply-connected.*

Proof. Fix a point $p \in M$. In view of Lemma 6.5.1, we know that $\exp_p : T_pM \to M$ is a local diffeomorphism. This being so, we may endow T_pM with the pull-back metric $\tilde{g} = \exp_p^* g$. Since a local isometry maps geodesics to geodesics, the geodesics of $(T_pM, \tilde{g})$ through the origin 0_p are the straight lines, thus, defined on all of $\mathbf{R}$ due to the completeness of M. In view of Theorem 3.3.5(c), this implies that $(T_pM, \tilde{g})$ is complete. Now $\exp_p$ is a covering because of Proposition 3.3.8(b), and the last assertion in the statement is obvious. □

A complete simply-connected manifold of nonpositive sectional curvature is called a *Hadamard manifold*.

6.5.3 Corollary *Let M be a Hadamard manifold. Then, given p, $q \in M$, there is a unique geodesic joining p to q.*

Proof. Let γ be a geodesic joining p to q. Consider the diffeomorphism $\exp_p : T_pM \to M$. Then $\exp_p^{-1} \circ \gamma$ is the straight line in T_pM joining the origin and $\exp_p^{-1}(q)$, as in the proof of Theorem 6.5.2, and this proves the uniqueness of γ. □

In particular, the preceding corollary implies that the cut-locus of an arbitrary point in a Hadamard manifold is empty.

The Hadamard-Cartan theorem says that the universal covering manifold of a complete Riemannian manifold M of nonpositive sectional curvature is $\mathbf{R}^n$. Since $\mathbf{R}^n$ is contractible, the higher homotopy groups $\pi_i(M)$, where $i \geq 2$, are all trivial. Consequently, the topological information about M is contained in its fundamental group $\pi_1(M)$. In the sequel, we prove some classical results about the fundamental group of nonpositively curved manifolds. We start with a lemma.

6.5.4 Lemma *Let M be a Hadamard manifold. Then, for any point $p \in M$, the function $f_p : M \to \mathbf{R}$ given by $f_p(x) = \frac{1}{2}d(p,x)^2$ is smooth and strictly convex.*

Proof. Fix a point $p \in M$. Denote by $\gamma^x : [0,1] \to M$ the unique geodesic parametrized with constant speed joining p to x. Plainly, γ^x is minimizing, so

$$f_p(x) = \frac{1}{2}L(\gamma^x)^2 = E(\gamma^x) = \frac{1}{2}||\gamma^{x\prime}(0)||^2 = \frac{1}{2}||\exp_p^{-1}(x)||^2,$$

showing that f_p is smooth.

Next, let η be a geodesic; we intend to verify that $f \circ \eta$ is strictly convex. For that purpose, we set $\gamma_t = \gamma^{\eta(t)}$ and invoke the second variation formula (5.3.9) to write:

$$\begin{aligned} \frac{d^2}{dt^2}\Big|_{t=0}(f_p \circ \eta)(t) &= \frac{d^2}{dt^2}\Big|_{t=0} E(\gamma_t) \\ &= \langle \overline{\nabla}_{\frac{\partial}{\partial t}} \frac{\bar{\partial}}{\partial t}\Big|_{t=0}, \gamma' \rangle \Big|_0^1 + \int_0^1 ||Y'||^2 + \langle R(\gamma', Y)\gamma', Y\rangle \, ds. \end{aligned} \tag{6.5.5}$$

Since the variational vector field $Y = \frac{\bar{\partial}}{\partial t}|_{t=0}$ vanishes at $s = 0$ and $\overline{\nabla}_{\frac{\partial}{\partial t}} \frac{\bar{\partial}}{\partial t}\Big|_{\substack{s=1\\t=0}} = \eta''(0) = 0$, the first term in the sum is zero; the assumption on the curvature and the fact that Y is nonzero imply that the second term there is positive. We conclude that f is strictly convex. □

6.5.6 Remark We can get more refined information about the second derivatives of f_p. It immediately follows from the Cauchy-Schwarz inequality that a smooth function $f : [0,1] \to \mathbf{R}$ with $f(0) = 0$ must satisfy the inequality $\int_0^1 (f')^2 \, ds \geq f(1)^2$. Retaining the notation in the proof of

Lemma 6.5.4, we write $Y(s) = \sum_i a_i(s)E_i(s)$ for smooth functions $a_i : [0,1] \to \mathbf{R}$ and an orthonormal frame $\{E_i\}$ of parallel vectors along γ_0. Then

$$\begin{aligned} \int_0^1 ||Y'||^2\, ds &= \sum_i \int_0^1 (a_i)'^2\, ds \\ &\geq \sum_i a_i(1)^2 \\ &= ||Y(1)||^2 \\ &= ||\eta'(0)||^2. \end{aligned}$$

Together with (6.5.5), this shows that (see exercise 13 in chapter 4)

$$\text{Hess}(f_p) \geq g$$

at every point of M, as bilinear symmetric forms. ⋆

Lemma 6.5.4 allows one to generalize the notion of center of mass of a finite set of points in Euclidean space to the context of Hadamard manifolds. For that purpose, two remarks are in order. First, we note that a non-negative strictly convex proper function has a unique minimum. In fact, because of properness, there must a minimum. If there were two minima, the function would be strictly convex when restricted to a geodesic joining the two minima, and this would imply that the function has smaller values on the interior of this segment than at the endpoints, contradicting the fact that the endpoints are minima. The second remark is that the maximum of any number of strictly convex functions is still strictly convex, as one sees easily. Now, given a finite set of points $p_1, \ldots, p_k$ in a Hadamard manifold, the *center of mass* of the set $\{p_1, \ldots, p_k\}$ is defined to be the uniquely defined minimum of the non-negative strictly convex proper function

$$x \mapsto \max\{f_{p_1}(x), \ldots, f_{p_k}(x)\}.$$

6.5.7 Theorem (Cartan) *Let M be a Hadamard manifold. Then any isometry of finite order of M has a fixed point.*

Proof. Let φ be an isometry of M of order $k \geq 1$. For an arbitrary point $p \in M$, set q to be the center of mass of the finite set $\{p, \varphi(p), \ldots, \varphi^{k-1}(p)\}$. This means that q is the unique minimum of the function

$$f(x) = \max\{f_p(x), f_{\varphi(p)}(x), \ldots, f_{\varphi^{k-1}(p)}(x)\}.$$

Since $\varphi^k(p) = p$ and φ is distance-preserving,

$$\begin{aligned} f(\varphi(q)) &= \frac{1}{2} \max\left\{d(p,\varphi(q))^2, d(\varphi(p),\varphi(q))^2, \ldots, d(\varphi^{k-1}(p),\varphi(q))^2\right\} \\ &= \frac{1}{2} \max\left\{d(\varphi^{k-1}(p),q)^2, d(p,q)^2, \ldots, d(\varphi^{k-2}(p),q)^2\right\} \\ &= f(q), \end{aligned}$$

which shows that also $\varphi(q)$ is a minimum of f. Hence, $\varphi(q) = q$. □

6.5.8 Corollary *Let M be a complete Riemannian manifold of nonpositive sectional curvature. Then the fundamental group of M is torsion-free.*

Proof. The Riemannian universal covering $\tilde{M}$ of M is a Hadamard manifold, and the elements of $\pi_1(M)$ act on $\tilde{M}$ as deck transformations, thus, without fixed points; Theorem 6.5.7 implies that they cannot have finite order. $\square$

Before proving the next theorem, we recall some facts about the relation between the fundamental group $\pi_1(M,p)$ and the set of free homotopy classes of loops, which we denote by $[S^1, M]$, for a connected manifold M and $p \in M$.

6.5.9 Lemma *The 'forgetful' map $\mathcal{F} : \pi_1(M,p) \to [S^1, M]$, which is obtained by ignoring basepoints, sets up a one-to-one correspondence between $[S^1, M]$ and the set of conjugacy classes in $\pi_1(M,p)$.*

Proof. First we remark that $\mathcal{F}$ is onto. In fact, let $\zeta_1 : [0,1] \to M$ be a loop in M, with $\zeta_1(0) = \zeta_1(1) = q$, representing a class in $[S^1, M]$. Since M is arcwise connected, there is a continuous path c joining p to q. Then $\zeta_t := c|_{[t,1]} \cdot \zeta \cdot (c|_{[t,1]})^{-1}$ is a continuous homotopy between ζ_0 and ζ_1, and ζ_0 lies in the image of $\mathcal{F}$.

Next, if γ, η are loops based at p then $\mathcal{F}[\eta \cdot \gamma \cdot \eta^{-1}] = \mathcal{F}[\eta] \cdot \mathcal{F}[\gamma] \cdot \mathcal{F}[\eta^{-1}] = \mathcal{F}[\eta^{-1}] \cdot \mathcal{F}[\eta] \cdot \mathcal{F}[\gamma] = \mathcal{F}[\gamma]$, where for the second equality we cyclically permute the order of concatenation by changing the basepoint. This proves that $\mathcal{F}$ is constant on conjugacy classes.

Conversely, let γ_0, $\gamma_1 : [0,1] \to M$ be loops based at p with $\mathcal{F}[\gamma_0] = \mathcal{F}[\gamma_1]$. This means there is a homotopy γ_t between those curves without necessarily preserving basepoints. The curve $c(t) = \gamma_t(0) = \gamma_t(1)$ traces out the path taken by the basepoints and thus is a loop. Now the concatenation $\tilde{\gamma}_t := c|_{[0,t]} \cdot \gamma_t \cdot (c|_{[0,t]})^{-1}$ is a homotopy from γ_0 to $c \cdot \gamma_1 \cdot c^{-1}$ preserving basepoints. $\square$

6.5.10 Lemma *Let γ, η be loops in M based at p, q, respectively. Then the classes $[\gamma] = [\eta]$ in $[S^1, M]$ if and only if $[\gamma] \in \pi_1(M,p)$ and $[\eta] \in \pi_1(M,q)$ act by the same deck transformation on the universal cover $\tilde{M}$.*

Proof. Let ζ be a curve joining p to q. Then $\zeta \cdot \eta \cdot \zeta^{-1}$ is in the same free homotopy class as η. Using Lemma 6.5.9, by concatenating ζ with a loop at p, we may assume that $[\zeta \cdot \eta \cdot \zeta^{-1}] = [\gamma]$ in $\pi_1(M,p)$. The desired result follows from the standard relation between the fundamental group and deck transformations. $\square$

6.5.11 Theorem (Preissmann) *Let M be a compact Riemannian manifold of negative sectional curvature. Then every nontrivial Abelian subgroup of its fundamental group is infinite cyclic.*

Proof. We can assume that M is not simply-connected. Let $\tilde{M}$ be the Riemannian universal covering of M, and let $\varphi \in \pi_1(M)$ an element different from the identity which we view as an isometry of $\tilde{M}$. Recall that φ acts on $\tilde{M}$ without fixed points. The fundamental remark is that the *displacement function* $f : \tilde{M} \to \mathbf{R}$ given by $f(x) = d(x, \varphi(x))$ is smooth and convex. For the purpose of proving this claim, consider the function $\Phi : TM \to M \times M$, given by $\Phi(v) = (x, \exp_x(v))$ for $v \in T_xM$, that was introduced in Lemma 2.4.6. Since $\tilde{M}$ is a Hadamard manifold, we easily see that Φ is well defined and a global diffeomorphism. Now $d : \tilde{M} \times \tilde{M} \setminus \Delta_{\tilde{M}} \to \mathbf{R}$ is given by $d(x,y) = g_x(\Phi^{-1}(x,y), \Phi^{-1}(x,y))^{1/2}$, so it is also smooth; here $\Delta_{\tilde{M}}$ denotes the diagonal of $\tilde{M}$. This proves that f is smooth. In order to prove the convexity of f, we resort to the second variation formula of the length given in exercise 1 of chapter 5. Let η be a geodesic; similarly to in (6.5.5), we can write

$$\frac{d^2}{dt^2}\Big|_{t=0}(f \circ \eta)(t) = \int_0^1 ||Y_\perp'||^2 + \langle R(\gamma', Y_\perp)\gamma', Y_\perp \rangle \, ds \geq 0, \tag{6.5.12}$$

where γ_t is the geodesic joining $\eta(t)$ to $\varphi(\eta(t))$, Y is the variational vector field along γ_0 and $Y_\perp$ denotes its normal component, and we have used that $\overline{\nabla}_{\frac{\partial}{\partial t}}\frac{\bar{\partial}}{\partial t}\big|_{t=0}$ is equal to $\eta''(0) = 0$ and $(\varphi\circ\eta)''(0) = 0$ for $s = 0$ and 1, respectively. Although f is not strictly convex, we can derive more refined information from formula (6.5.12). Since $\tilde{M}$ has negative curvature, the equality holds in (6.5.12) if and only if Y is a constant multiple of γ', so at any given point $x \in \tilde{M}$, f is stricly convex in any direction different from the direction of the geodesic joining x to $\varphi(x)$.

Next, we introduce a definition. An *axis* of φ is a geodesic of $\tilde{M}$ that is invariant under φ. Note that φ cannot reverse the orientation of an axis γ for otherwise the midpoint of the geodesic segment between $\gamma(t)$ and $\varphi(\gamma(t))$ would be a fixed point of φ for any $t \in \mathbf{R}$. Hence the restriction of φ to γ must be translation along it:

$$\varphi(\gamma(t)) = \gamma(t + t_0)$$

for some $t_0 \in \mathbf{R}$ and all $t \in \mathbf{R}$. The number t_0 will be called the *period* of φ along the axis γ. For later reference, we also note that

$$f(\varphi(x)) = d(\varphi(x), \varphi^2(x)) = d(x, \varphi(x)) = f(x)$$

for every $x \in \tilde{M}$.

Now we give three important properties of axes. The first one is that f is constant along an axis γ of φ. Indeed,

$$f(\gamma(t + t_0)) = f(\varphi(\gamma(t))) = f(\gamma(t))$$

for all $t \in \mathbf{R}$, where t_0 is the period of γ. It follows that $f \circ \gamma$ is convex and periodic, and it is easy to see that such a function must be constant. The second one is that an axis of φ is a set of minima of f. This follows immediately from the formula of the first variation of length and (6.5.12). The last one is that if f is constant on a geodesic segment $\overline{xy}$ for points $x \neq y$, then the supporting geodesic γ of that segment is an axis of φ. Indeed, f is not stricly convex along $\overline{xy}$, so γ must coincide with the geodesic joining x and $\varphi(x)$. It follows that $\varphi(x)$ lies in the image of γ. Similarly, $\varphi(y)$ lies in the image of γ. Since a geodesic in $\tilde{M}$ is uniquely defined by two points on it, γ must be an axis of φ.

The next step is to prove that φ admits one and only one axis, up to reparametrization and reorientation. Note that the value f at a point $x \in \tilde{M}$ is the length of the unique geodesic in $\tilde{M}$ joining x to $\varphi(x)$. Such geodesics project to geodesics in M all lying in the same free homotopy class of loops in M, independent of the point x, according to Lemma 6.5.10. Since M is compact, f admits a global minimum $p \in \tilde{M}$ by Lemma 6.3.1. Since $f(\varphi(p)) = f(p)$, we have that $\varphi(p)$ is also a global minimum. By convexity, f is constant along the geodesic segment joining p and $\varphi(p)$; let γ be the unit speed geodesic that supports this segment. By the above, γ is an axis of φ. Now the points in the image of γ comprise a set of minima at each point of which f is strictly convex in any direction different from γ. It follows that there cannot be another axis.

Finally, suppose that H is an Abelian subgroup of $\pi_1(M)$, and that φ belongs to H and has γ as an axis as above. Since the elements of H commute with φ, they map γ to a geodesic which is invariant under φ; by the above uniqueness result, γ is an axis for all the elements of H. Consider now the period map $H \to \mathbf{R}$. This map is clearly an injective homomorphism, thus its image is a subgroup of $\mathbf{R}$ isomorphic to H. It is not difficult to see that every subgroup of $\mathbf{R}$ is either infinite cyclic or dense. Since the orbits of H on $\tilde{M}$ are discrete, H must be infinite cyclic. □

6.5.13 Corollary *No compact nontrivial product manifold $M \times N$ admits a metric with negative sectional curvature.*

Proof. Suppose, on the contrary, that $M \times N$ supports a metric of negative sectional curvature. By the Hadamard-Cartan theorem 6.5.2, its universal covering, which is the product of the universal coverings $\tilde{M}$ of M and $\tilde{N}$ of N, is contractible. Since a compact manifold can never be contractible (unless it is a point), neither $\tilde{M}$ nor $\tilde{N}$ is compact. In particular, neither M nor N is simply-connected. Now $\pi_1(M)$ and $\pi_1(N)$ are both non-trivial, and as subgroups of $\pi_1(M \times N)$, each of its elements have infinite order by Corollary 6.5.8. We deduce that $\pi_1(M)$ and $\pi_1(N)$ contain infinite cyclic groups H and K, respectively. But then $H \times K$ is a non-trivial Abelian subgroup of $\pi_1(M \times N)$ which is not cyclic, contradicting Preissmann's theorem. This proves the corollary. □

6.5.14 Remark An isometry φ of a Hadamard manifold $\tilde{M}$ can be of three types. Let f be the displacement function associated to φ as in Preissmann's theorem 6.5.11. Then φ is said to be:

a. *elliptic* if f attains the value zero (i.e. φ admits a fixed point);

b. *hyperbolic* if f attains a positive minimum;

c. *parabolic* if f attains no minimum.

The argument in Preissmann's theorem proves that a hyperbolic isometry of a Hadamard manifold admits an axis (which is unique in the case in which the curvature of $\tilde{M}$ is negative).

6.6 Rauch's theorem

In this section we present a version of Rauch's theorem, which is another example of comparison theorem in Riemannian geometry, and derive as an application the existence of convex neighborhoods in Riemannian manifolds.

6.6.1 Theorem (Rauch) *Let $\gamma : [0, \ell] \to M$ be a unit speed geodesic in a Riemannian manifold M and assume that the sectional curvatures of M along γ are bounded above by a real constant κ. If Y is a Jacobi field along γ which is always orthogonal to γ', then the function $||Y||$ along γ satisfies the differential inequality*

$$||Y||'' + \kappa ||Y|| \geq 0 \qquad (6.6.2)$$

on the complement of the zero set of Y on $(0, \ell)$.

Moreover, if ψ denotes the solution on $[0, \ell]$ of the differential equation

$$\psi'' + \kappa\psi = 0, \quad \psi(0) = ||Y||(0), \quad \psi'(0) = ||Y||'(0),$$

and ψ does not vanish on $(0, \ell)$, then Y does not vanish on $(0, \ell)$ and

$$\left(\frac{||Y||}{\psi}\right)' \geq 0 \quad \text{and} \quad ||Y|| \geq \psi \qquad (6.6.3)$$

on $(0, \ell)$.

Finally, the first inequality in (6.6.3) is an equality at s_0 for some $s_0 \in (0, \ell)$ if and only if the sectional curvatures $K(\gamma', Y) = \kappa$ along $[0, s_0]$ and there exists a parallel unit vector field E along γ for which

$$Y(t) = \psi(t)E(t)$$

along $[0, s_0]$.

Proof. We differentiate $||Y||$ twice along γ to obtain

$$||Y||' = \frac{\langle Y', Y\rangle}{||Y||}$$

and

$$\begin{aligned} ||Y||'' &= \frac{\left(||Y'||^2 + \langle Y'', Y\rangle\right)||Y|| - \langle Y', Y\rangle^2/||Y||}{||Y||^2} \\ &= \frac{||Y'||^2||Y||^2 - \langle Y', Y\rangle^2}{||Y||^3} + \frac{1}{||Y||}\langle R(\gamma', Y)\gamma', Y\rangle \\ &\geq -\kappa||Y||, \end{aligned}$$

where we have used the Jacobi equation, the Cauchy-Schwarz inequality and the assumption that the sectional curvature of the plane spanned by γ', Y is bounded above by κ, proving the differential inequality.

Moreover, if ψ is as in the statement, then

$$\left(\frac{||Y||}{\psi}\right)' = \frac{||Y||'\psi - ||Y||\psi'}{\psi^2},$$

where the numerator satisfies

$$(||Y||'\psi - ||Y||\psi')(0) = 0$$

by the assumptions, and

$$\left(||Y||'\psi - ||Y||\psi'\right)' = ||Y||''\psi - ||Y||\psi'' \geq 0$$

on $(0, s_0)$ by the differential inequality, where $s_0 > 0$ is the first parameter value where $Y(s_0) = 0$. It follows that the numerator is also non-negative, proving that $(||Y||/\psi)' \geq 0$ on $[0, s_0]$. Since $\lim_{s\to 0^+} \frac{||Y(s)||}{\psi(s)} = 1$, this implies that $||Y|| \geq \psi$ on $[0, s_0]$. Finally, the assumption that ψ does not vanish on $(0, \ell)$ shows that $s_0 \geq \ell$.

If we have equality in the first equation in (6.6.3) for some $s_0 \in (0, \ell)$, then we have equality on all of $(0, s_0]$, which implies $||Y|| = \psi$ on all of $[0, s_0]$. Write $Y = \psi E$ where $||E|| = 1$ along γ. Then $Y' = \psi' E + \psi E'$. We have equality in (6.6.2), which implies $K(\gamma', Y) = \kappa$ and also equality in the Cauchy-Schwarz inequality above, meaning that Y and Y' are linearly dependent at every point of $(0, s_0]$; hence E is parallel along $\gamma|_{[0,s_0]}$. □

The following corollary of Rauch's theorem is attributed to M. Morse (1930) and I. J. Schönberg (1932), and is a strengthening of Lemma 6.5.1.

6.6.4 Corollary *Let $\gamma : [0, \ell] \to M$ be a unit speed geodesic in a Riemannian manifold M and assume that the sectional curvatures of M along γ are bounded above by a positive real constant κ. Then the first conjugate point of $\gamma(0)$ along γ can only occur at $s \geq \pi/\sqrt{\kappa}$.*

Proof. Let Y be a Jacobi field along γ with $Y(0) = 0$; by rescaling, we may assume $||Y||'(0) = 1$. In Rauch's theorem, we have $\psi(s) = \frac{\sin(\sqrt{\kappa}s)}{\sqrt{\kappa}}$, and $||Y(s)|| \geq \frac{\sin(\sqrt{\kappa}s)}{\sqrt{\kappa}} > 0$ for $s < \pi/\sqrt{\kappa}$. □

6.6.5 Lemma *Let p be a point in a Riemannian manifold and choose a sufficiently small $r > 0$ such that $B(p, r)$ is a normal neighborhood of p and $r < \frac{\pi}{2\sqrt{\kappa}}$, where κ is the supremum of sectional curvatures of M at points in a given compact neighborhood of p, and we interpret $\frac{\pi}{2\sqrt{\kappa}}$ as $+\infty$ in case $\kappa \leq 0$. If $\eta : [0, 1] \to B(p, r)$ is a geodesic segment, then the function $f(t) = d(p, \eta(t))$ has at most one critical point for $t \in (0, 1)$, and such a critical point must be a point of minimum.*

Proof. It suffices to prove that any critical point $t_0 \in (0,1)$ of f is a point of minimum. Construct a smooth variation through geodesics $\{\gamma_t\}$ where $\gamma_t : [0,\ell] \to B(p,r)$ is the unique constant speed geodesic joining p to $\eta(t)$ and γ_{t_0} has unit speed. Note that $\ell < r < \frac{\pi}{2\sqrt{\kappa}}$. The variational vector field Y is a Jacobi field orthogonal to γ' at the endpoints, and thus everywhere. By the second variation formula of length (exercise 1 of chapter 5),

$$\frac{d^2}{dt^2}\Big|_{t=t_0} L(\gamma_t) = \langle \nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial t}\Big|_{t=0}, \gamma' \rangle\Big|_0^\ell + \int_0^\ell ||Y'||^2 + \langle R(\gamma', Y)\gamma', Y\rangle\, ds.$$

Since $\gamma_t(0) = p$ for all t, and $t \mapsto \gamma_t(\ell) = \eta(t)$ is a geodesic, the first term on the right-hand side vanishes. In case $\kappa \leq 0$, this already shows that $\frac{d^2}{dt^2}|_{t=t_0} L(\gamma_t) > 0$ and hence t_0 is a point of minimum. Otherwise $\kappa > 0$ and, using $||Y'||^2 = \langle Y, Y'\rangle' - \langle Y, Y''\rangle$ and the Jacobi equation, we can write

$$\frac{d^2}{dt^2}\Big|_{t=t_0} L(\gamma_t) = \langle Y'(\ell), Y(\ell)\rangle.$$

By Rauch's theorem 6.6.1,

$$\frac{||Y||'}{||Y||} \geq \frac{\psi'}{\psi}$$

on $(0,\ell)$, where $\psi(s) = \sin(\sqrt{\kappa}s)\frac{||Y||'(0)}{\sqrt{\kappa}}$. It follows that

$$\langle Y'(\ell), Y(\ell)\rangle = ||Y||'(\ell)||Y(\ell)|| \geq \frac{\psi'(\ell)}{\psi(\ell)}||Y(\ell)||^2 = \sqrt{\kappa}\cot(\sqrt{\kappa}\ell)||Y(\ell)||^2 > 0,$$

which proves that t_0 is a point of minimum. □

In Proposition 2.4.7, we proved the existence of a totally normal neighborhood U of any point in a Riemannian manifold, namely, any two points in U can be connected by a unique minimizing geodesic. We next show that U can be chosen so that the minimizing geodesic lies entirely in U.

A subset C of a Riemannian manifold M is called *strongly convex* if every two points p, q lying in the topological closure $\bar{C}$ can be connected by a unique minimizing geodesic $\eta : [0,1] \to M$ such that $\eta(0,1) \subset C$. J. H. C. Whitehead proved in 1932 that any point in any Riemannian manifold is the center of a sufficiently small open metric ball which is strongly convex. Recall that the injectivity radius inj as a function on a Riemannian manifold, namely, the distance of a point to its cut locus, is a continuous function.

6.6.6 Theorem (Whitehead) *Let p be a point in a Riemannian manifold. Fix a compact neighborhood K of p in M, let ι denote the infimum of* inj *over K, and let κ denote the supremum of sectional curvatures at points in K. If $r < \frac{1}{2}\min\{\frac{\pi}{\sqrt{\kappa}}, \iota\}$ and $B(p,r) \subset K$, then $B(p,r)$ is strongly convex; here we interpret $\frac{\pi}{\sqrt{\kappa}}$ as $+\infty$ in case $\kappa \leq 0$.*

Proof. Let q, $q' \in \bar{B}(p,r)$. Then $d(q,q') \leq 2r < \iota$, so there is a unique minimizing geodesic $\gamma; [0,1] \to M$ connecting q to q' and γ depends continuously on q, q'. Choose $\epsilon > 0$ such that $r + \epsilon < \frac{1}{2}\min\{\frac{\pi}{\kappa}, \iota\}$, and put

$$V_{r+\epsilon} = \{\, (q,q') \in \bar{B}(p,r) \times \bar{B}(p,r) \mid \gamma(0,1) \subset B(p, r+\epsilon)\,\}.$$

This set is clearly non-empty, and also open in $\bar{B}(p,r) \times \bar{B}(p,r)$ since γ depends continuously on its endpoints. Owing to Lemma 6.6.5,

$$\gamma(0,1) \subset B(p,r) \quad \text{for all} \quad (q,q') \in V_{r+\epsilon}. \tag{6.6.7}$$

Again by continuous dependence of γ on its endpoints, we have that $(q, q') \in \bar{V}_{r+\epsilon}$ implies $\gamma[0,1] \subset \bar{B}(p,r) \subset B(p, r+\epsilon)$, therefore $\bar{V}_{p+\epsilon} \subset V_{r+\epsilon}$, which means that $V_{r+\epsilon}$ is closed. By connectedness, $V_{r+\epsilon} = \bar{B}(p,r) \times \bar{B}(p,r)$, and we finish the proof by referring to (6.6.7). □

The *convexity radius at p* is the supremum (which may be $+\infty$) of all $r \in \mathbb{R}$ such that, for all $\eta < r$, the geodesic ball $B(p, \eta)$ is strongly convex. The *convexity radius* of M is the infimum of convexity radii at all points of M. For instance, the convexity radius of the sphere is $\pi/2$.

6.7 Additional notes

§1 The Gauss-Lobatchevsky-Bolyai discovery of hyperbolic geometry in the early nineteenth century finally pointed out the impossibility of proving Euclid's fifth postulate from the other postulates of Euclidean geometry. In 1868, Beltrami proved the consistency of hyperbolic geometry by realizing it as the intrinsic geometry of a well known surface in Euclidean 3-space — the so-called pseudo-sphere — which has constant negative curvature. In his *Habilitationsvortrag* of 1854 in which Riemann laid the foundations of Riemannian geometry were also exhibited examples of metrics of arbitrary constant curvature. Based on Riemann's ideas, Beltrami published another article in 1869 in which he discussed spaces of constant curvature in arbitrary dimensions. In this way, the non-Euclidean geometries were for the first time incorporated into the realm of Riemannian geometry. In 1890, Klein drew attention to Clifford's 1873 discovery of a 2-torus — nowadays known as the *Clifford torus* — sitting in S^3 with constant zero curvature and formulated the problem of classifying Riemannian manifolds of arbitrary constant curvature in arbitrary dimensions. The problem, referred to as the *Clifford-Klein space forms problem*, was extensively studied by Killing in an article in 1891 and a book in 1893, and then again by Heinz Hopf in 1925 culminating in Theorem 6.2.2.

§2 The argument in the proof of the Hadamard-Cartan theorem 6.5.2 shows that if there is a point in a simply-connected Riemannian manifold possessing no conjugate points, then the manifold is diffeomorphic to Euclidean space. Eberhard Hopf [Hop48] proved that a compact Riemannian manifold M without conjugate points satisfies the inequality

$$\int_M \mathrm{scal} \leq 0$$

where the integral is taken with respect to the canonical Riemannian measure (exercise 12 of chapter 4), and the equality holds if and only if M is flat. In the 2-dimensional case, the left-hand side equals 2π times the Euler characteristic of M by the Gauss-Bonnet theorem. It follows from E. Hopf's result that a metric without conjugate points on T^2 must be flat. It was a long standing conjecture that the same result should be also valid for the higher dimensional tori. In 1994, Burago and Ivanov [BI94] finally settled the conjecture in the positive sense.

§3 Techniques from geometric analysis have been proved to be very powerful in dealing with problems involving curvature in Riemannian manifolds. We would like to mention two spectacular instances of this fact. In 1960, Yamabe [Yam60] tried to deform conformally a given Riemannian metric g on a manifold M into a metric $f \cdot g$ of constant scalar curvature, where f is an unknown positive smooth function on M. If $n = \dim M = 2$, this is a classical result and amounts to showing that M admits isothermal coordinates [Jos06], so he was dealing with the case $n \geq 3$. There was a problem with Yamabe's arguments, though, and the question became the *Yamabe problem*. In order to find f, one needs to solve the nonlinear partial differential equation

$$\Delta f + \frac{n-2}{4(n-1)} \mathrm{scal}(M, g) = f^{\frac{n+2}{n-2}}.$$

This is an extremely difficult question in analysis because the exponent of f is exactly the "critical exponent" in regard to which the standard Sobolev embedding theorems do not apply. The problem was eventually solved through the work of of Aubin [Aub76] and Schoen [Sch84]. Thanks to contributions by other mathematicians, the Yamabe problem is today almost completely understood and it is known that the set of metrics of constant scalar curvature in a given conformal class of metrics is an infinite-dimensional space if $n > 2$. See [Aub98] for these results in book form.

Deformation techniques like that concerning the Yamabe problem are used to prove the existence of several objects in geometry. An interesting approach is to consider deformations on the level of the space of Riemannian metrics on a given smooth manifold M. For instance, Hamilton [Ham82] introduced the following normalized *Ricci flow* equation in the space of Riemannian metrics on a compact n-dimensional manifold M:

$$\frac{d}{dt}g(t) = -2\mathrm{Ric}(g(t)) + 2\frac{\tau}{n}g(t),$$

where $\mathrm{Ric}(g(t))$ denotes the Ricci curvature of the metric $g(t)$, and τ denotes the integral of the scalar curvature of $g(t)$. The fixed points of this equation are the metrics of constant Ricci curvature. One considers t as time and studies the equation as an initial value problem for a fixed Riemannian metric $g_0 = g(0)$ on M. Hamilton proved that if $n = 3$ and the Ricci curvature of g_0 is positive, then the Ricci flow converges smoothly to a metric of constant Ricci curvature. In particular, the manifold is diffeomorphic to a spherical space form. At that time, this was a very interesting application of Riemannian geometry to provide a partial answer to a long-standing open problem in topology, the so called *Poincaré conjecture*: Is every simply-connected compact 3-dimensional manifold homeomorphic to S^3? The difficulty in using Hamilton's method to prove the full Poincaré conjecture was that if one removes the assumption that $\mathrm{Ric}(g_0) > 0$, then the Ricci flow develops finite-time singularities that impede the convergence to a nice metric, and those singularities were not completely understood. As it turns out, Perelman was able to overcome those analytic difficulties. He extended Hamilton's results and in particular proved the full Poincaré conjecture (see e.g. [MT06]).

§4 A famous, open conjecture of Heinz Hopf asserts that $S^2 \times S^2$ does not admit a metric of positive sectional curvature. Indeed, known examples of simply-connected compact manifolds with positive sectional curvature are relatively rare (owing to the Bonnet-Myers theorem 6.4.1, the non-simply-connected examples are quotients of the simply-connected ones by finite subgroups of isometries). The standard examples are the compact rank one symmetric spaces (see Add. notes ? of chapter ?). Apart from these, the homogeneous examples have been classified by Wallach [Wal72] in the odd-dimensional case and by Bérard-Bergery [BB76] in the even dimensional case. These examples occur only in dimensions 6, 7, 12, 13 and 24, and are due to Berger, Wallach and Allof-Wallach. The only other examples known are given by *biquotients* $G//H$. Here G is a Lie group equipped with a bi-invariant metric and H is subgroup of $G\times G$ acting on G by $(h_1, h_2)\cdot g = h_1 g h_2^{-1}$. This action is always proper and isometric, and if it is also free, then the quotient space is a manifold denoted by $G//H$. In this case, there is a unique metric on $G//H$ making the projection $G \to G//H$ into a Riemannian submersion and it follows from Proposition 4.5.8 that $G//H$ has always non-negative curvature. More generally, one can also construct bi-quotients by considering left-invariant metrics on G more general than the bi-invariant ones. It turns out that the only known examples of positively curved biquotients occur in dimensions 6, 7 and 13, and these are due to Eschenburg and Bazaikin. There is no general classification of positively curved biquotients. See [Zil07] for a recent survey on these results and related ones.

6.8 Exercises

1 Some definitions: a Riemannian manifold M is called *locally homogeneous* is any two points admit isometric neighborhoods. The *local isotropy group* of M at a point p is the group of germs of isometries defined on connected neighborhoods of p; note that this group is well defined in view of exercise 15 of chapter 3. Finally, M is called *locally* 3*-point homogeneous* if for any two points p_0, p_0' there exist connected neighborhoods U, U' of p_0, p_0', resp., such that given two triples (p,q,r), (p',q',r') of points in U, U', resp., with $d(p,q) = d(p',q')$, $d(q,r) = d(q',r')$, $d(r,p) = d(r',p')$, there exists a distance-preserving map $f : U \to M$ that maps the first triple to the second one.

Let M be a complete Riemannian manifold of dimension n. Prove that the following assertions are equivalent:

a. M has constant sectional curvature.
b. M is locally homogeneous, and its local isotropy group at any point is isomorphic to $\mathbf{O}(n)$.
c. M is locally 3-point homogeneous.

2 Prove that an odd-dimensional compact Riemannian manifold of positive sectional curvature is orientable.

3 Let M be a complete Riemannian manifold of nonpositive curvature. Prove that each homotopy class of curves with given endpoints in M contains a unique geodesic.

4 Consider the ball model $\mathbf{D}^n$ of $\mathbf{R}H^n$ and let φ be an isometry of $\mathbf{R}H^n$.

a. Prove that φ uniquely extends to a homeomorphism of the closed ball $\overline{\mathbf{D}^n}$. (Hint: Use exercise 4 of chapter 3.)
b. Prove that φ is hyperbolic if and only if its extension to $\overline{\mathbf{D}^n}$ admits exactly two fixed points and those lie in the boundary S^{n-1}.
c. Prove that φ is parabolic if and only if its extension to $\overline{\mathbf{D}^n}$ admits exactly one fixed point and that lies in the boundary S^{n-1}.

5 In the notation of exercise 5 of chapter 1, assume that the isometry T of the upper-half-plane is not the identity and prove that it is hyperbolic, elliptic or parabolic according to whether $(a-d)^2 + 4bc$ is positive, negative or zero, respectively.

6 Let G be an Abelian subgroup of the fundamental group of a spherical space form M. Prove that G is cyclic.

7 An isometry φ of a Riemannian manifold M is called a *Clifford translation* if the associated displacement function $x \mapsto d(x, \varphi(x))$ is constant. Prove that:

a. The Clifford translations for $\mathbf{R}^n$ are just the ordinary translations.
b. The only Clifford translation of $\mathbf{R}H^n$ is the identity transformation.
c. A linear transformation $A \in \mathbf{O}(n+1)$ is a Clifford translation of S^n if and only if there is a unimodular complex number λ such that half the eigenvalues of A are λ and the other half are $\bar{\lambda}$.

8 Let M be a Hadamard manifold. Prove that an isometry φ of M is a Clifford translation (cf. exercise 7) if and only if the vector field X on M given by $\exp_p(X_p) = \varphi(p)$ is parallel.

9 Extend Preissmann's theorem 6.5.11 to show that every solvable subgroup of the fundamental group of a compact Riemannian manifold of negative curvature must be infinite cyclic.

10 In this exercise, we prove that a compact homogeneous Riemannian manifold M whose Ricci tensor is negative semidefinite everywhere is isometric to a flat torus.

a. Use exercise 9 of chapter 5 to show that the identity component of the isometry group of M is Abelian.

b. Check that M can be identified with an n-torus equipped with a left-invariant Riemannian metric.

c. Show that an n-torus equipped with a left-invariant Riemannian metric admits a global parallel orthonormal frame and hence is flat.

11 A Riemannian manifold M is called *locally symmetric* if every point $p \in M$ admits a normal neighborhood V and an isometry $\varphi : V \to V$ such that $\varphi(p) = p$ and $d\varphi_p = -\mathrm{id}$.

a. Show that space forms and Lie groups with bi-invariant metrics are locally symmetric. (Hint: for the second example, use group inversion.)

b. Prove that the curvature tensor of a locally symmetric manifold is parallel. (Hint: Use the version of equation (4.2.6) for ∇R.)

12 Let M be a Riemannian manifold with curvature tensor R.

a. Prove that R is parallel if and only if for every smooth curve γ in M and parallel vector fields X, Y, Z, W along γ we have that $\langle R(X,Y)Z, W\rangle$ is constant.

b. Prove that if R is parallel then the Jacobi equation along a geodesic has constant coefficients in a suitable basis.

13 In this exercise, we prove the converse of the result of exercise 11(a).

a. Let M and $\tilde{M}$ be a Riemannian manifolds with parallel curvature tensors. Suppose there are points $p \in M$, $\tilde{p} \in \tilde{M}$ and a linear isometry $f : T_pM \to T_{\tilde{p}}\tilde{M}$ that takes any 2-plane in T_pM to a 2-plane in $T_{\tilde{p}}\tilde{M}$ with the same sectional curvature. Prove that there exists normal neighborhoods V, $\tilde{V}$ of p, $\tilde{p}$, resp., and an isometry $F : V \to \tilde{V}$ such that $F(p) = \tilde{p}$ and $dF_p = f$. (Hint: combine the idea in the proof of Theorem 6.2.1 with exercise 12(b)).

b. Prove that a Riemannian manifold with parallel curvature tensor is locally symmetric. (Hint: Apply part (a) to $M = \tilde{M}$ and $f = -\mathrm{id}$.)

Submanifold geometry

7.1 Introduction

In this chapter, we study the extrinsic geometry of Riemannian manifolds. Historically speaking, the field of Differential Geometry started with the study of curves and surfaces in $\mathbf{R}^3$, which originated as a development of the invention of infinitesimal calculus. Later investigations considered arbitrary dimensions and codimensions. Our discussion in this chapter is centered in submanifolds of spaces forms.

A most fundamental problem in submanifold geometry is to discover simple (sharp) relationships between intrinsic and extrinsic invariants of a submanifold. We begin this chapter by presenting the standard related results for submanifolds of space forms, with some basic applications. Then we turn to the Morse index theorem for submanifolds. This is a very important theorem that can be used to deduce information about the topology of the submanifold. Finally, we present a brief account of the theory of isoparametric submanifolds of space forms, which in some sense are the submanifolds with the simplest local invariants, and we refer to [BCO16, ch. 4] for a fuller account.

7.2 The fundamental equations of the theory of isometric immersions

The first goal is to introduce a number of invariants of the isometric immersion. From the point of view of submanifold geometry, it does not make sense to distinguish between two isometric immersions of M into $\overline{M}$ that differ by an ambient isometry. We call two isometric immersions $f:(M,g)\to(\overline{M},\overline{g})$ and $f':(M,g')\to(\overline{M},\overline{g})$ *congruent* if there exists an isometry φ of $\overline{M}$ such that $f'=\varphi\circ f$. In this case,

$$g' = f'^*\overline{g} = f^*\varphi^*\overline{g} = f^*\overline{g} = g.$$

Because of this, the induced metric is considered to be one of the basic invariants of an isometric immersion, and it is sometimes referred to as the *first fundamental form* of the immersion.

Due to the fact that our first considerations are local, we may assume that f is an embedding; for simplicity, we assume that f is the inclusion. In this case, for every point $p\in M$, the tangent space T_pM is a subspace of $T_p\overline{M}$ and the metric g_p is the restriction of $\overline{g}_p$. Consider the $\overline{g}$-ortohognal bundle decomposition

$$T\overline{M} = TM\oplus TM^\perp,$$

and denote by $(\cdot)^\top$ and $(\cdot)^\perp$ the respective projections. According to (2.8.2), the Levi-Civita connections ∇ and $\overline{\nabla}$ of M and $\overline{M}$, respectively, are related by

$$\nabla_XY = (\overline{\nabla}_{\overline{X}}\overline{Y})^\top,$$

where X and Y are vector fields on M and $\overline{X}$, $\overline{Y}$ are arbitrary extensions to vector fields on $\overline{M}$. The *second fundamental form* of the immersion f is the bilinear form $B : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM^\perp)$ given by

$$\begin{aligned} B(X,Y) &= \overline{\nabla}_{\overline{X}}\overline{Y} - \nabla_X Y \\ &= (\overline{\nabla}_{\overline{X}}\overline{Y})^\perp, \end{aligned}$$

where X, $Y \in \Gamma(TM)$ and $\overline{X}$, $\overline{Y} \in \Gamma(T\overline{M})$ are arbitrary local extensions of X, Y. In order to check that the definition of $B(X,Y)$ does not depend on the choice of local extensions, choose other ones $\overline{X}'$, $\overline{Y}' \in T\overline{M}$. Then

$$(\overline{\nabla}_{\overline{X}}\overline{Y} - \nabla_X Y) - (\overline{\nabla}_{\overline{X}'}\overline{Y}' - \nabla_X Y) = \overline{\nabla}_{\overline{X}-\overline{X}'}\overline{Y} - \nabla_{\overline{X}'}(\overline{Y}' - \overline{Y}).$$

Note that the right hand side vanishes at a point $p \in M$. Indeed, the first term is zero because $\overline{X}'_p = X_p = \overline{X}_p$, and the second term is zero because $\overline{Y}' = Y = \overline{Y}$ along a curve in M tangent to $\overline{X}'_p = X_p$ (cf. Remark 2.2.1). The second fundamental form is another one of the basic invariants of an isometric immersion. The orthogonal decomposition

$$\overline{\nabla}_X Y = \nabla_X Y + B(X,Y) \tag{7.2.1}$$

is called the *Gauss formula*.

We agree to retain the above notation and make some remarks about B. First, observe that $B(X,Y) = B(Y,X)$. This is because

$$\begin{aligned} B(X,Y) - B(Y,X) &= \overline{\nabla}_{\overline{X}}\overline{Y} - \nabla_X Y - \overline{\nabla}_{\overline{Y}}\overline{X} + \nabla_Y X \\ &= [\overline{X},\overline{Y}] - [X,Y] \\ &= 0 \end{aligned}$$

on M, where we have used the fact that $[\overline{X},\overline{Y}]$ is a local extension of $[X,Y]$. Next, note that it follows from the first defining condition of ∇ that B is $C^\infty(M)$-linear in the first argument; now it is a consequence of its symmetry that B is $C^\infty(M)$-linear also in the second argument. Therefore, for $p \in M$, $B(X,Y)_p$ depends only on X_p and Y_p. So there is a bilinear symmetric form

$$B_p : T_pM \times T_pM \to T_pM^\perp$$

given by $B_p(u,v) = B(U,V)_p$ where u, $v \in T_pM$, and U, V are arbitrary extensions of u, v to local vector fields on M. If $\xi \in T_pM^\perp$, the *Weingarten operator*, also called *shape operator* of the immersion f at ξ, is the self-adjoint linear endomorphism

$$A_\xi : T_pM \to T_pM,$$

given by

$$\langle A_\xi(u), v\rangle = \langle B(u,v), \xi\rangle,$$

where u, $v \in T_pM$. The eigenvalues of the Weingarten operator at ξ are called *principal curvatures* at ξ. Now the following lemma is proved by a simple computation.

7.2.2 Lemma *Let $\hat{\xi} \in \Gamma(TM^\perp)$ be a local extension of ξ to a normal vector field. Then*

$$A_\xi(u) = -(\overline{\nabla}_u\hat{\xi})^\top.$$

Proof. We have that

$$\langle A_\xi(u), v\rangle = \langle B(u,v), \xi\rangle = \langle \overline{\nabla}_{\overline{U}}\overline{V}, \hat{\xi}\rangle_p = -\langle \overline{V}, \overline{\nabla}_{\overline{U}}\hat{\xi}\rangle_p = -\langle v, \overline{\nabla}_u\hat{\xi}\rangle.$$

The result follows. □

The normalized trace of the second fundamental form

$$H = \frac{1}{n}\mathrm{tr}(B),$$

where $n = \dim M$, is called the *mean curvature vector* of M. Note that $n\langle \xi, H\rangle$ is the sum of the principal curvatures of M along ξ. A *minimal submanifold* is a submanifold with vanishing mean curvature. Minimal submanifolds are exactly the critical points of the volume functional (cf. exercise 12 of chapter 4) with respect to compactly supported variations. There is a vast literature devoted to them, especially in the case of minimal surfaces, which can be traced back at least to Euler and Lagrange. The minimal surface equation translates in coordinates to perhaps the best of all studied quasi-linear elliptic PDE, in terms of qualitative properties and explicit global solutions. For good introductions to minimal submanifolds, see [Law80, Sim83]. A classical reference to minimal surfaces is [Oss86]; a more recent one is [CM11].

Let us now turn to the last important invariant of an isometric immersion. Consider again the $\overline{g}$-orthogonal splitting $T\overline{M} = TM \oplus TM^\perp$. The bundle $TM^\perp \to TM$ is called the *normal bundle* of the isometric immersion. The connection in $T\overline{M}$ defines a connection $\nabla^\perp$ in $TM^\perp$, called the *normal connection* of the immersion, via the following formula,

$$\nabla^\perp_X\xi = (\overline{\nabla}_X\xi)^\perp,$$

where $\xi \in \Gamma(TM^\perp)$ and $X \in \Gamma(TM)$. It is a simple matter now to derive the *Weingarten formula*

$$\overline{\nabla}_X\xi = -A_\xi(X) + \nabla^\perp_X\xi; \tag{7.2.3}$$

indeed, we have

$$\nabla^\perp_X\xi = (\overline{\nabla}_X\xi)^\perp = \overline{\nabla}_X\xi - (\overline{\nabla}_X\xi)^\top = \overline{\nabla}_X\xi + A_\xi(X),$$

by Lemma 7.2.2, checking the equation. The normal connection is the third and last basic invariant of an isometric immersion that we wanted to mention.

Next, we want to state the fundamental equations involving the basic invariants of an isometric immersion. These are respectively called the *Gauss, Codazzi-Mainardi and Ricci equations.* At this juncture, we recall that the covariant derivative of the second fundamental form is given by

$$(\nabla^\perp_X B)(Y,Z) = \nabla^\perp_X(B(Y,Z)) - B(\nabla_X Y, Z) - B(Y, \nabla_X Z)$$

(cf. section 4.4), and the *normal curvature* of the immersion is given by

$$R^\perp(X,Y)\xi = \nabla^\perp_X\nabla^\perp_Y\xi - \nabla^\perp_Y\nabla^\perp_X\xi - \nabla^\perp_{[X,Y]}\xi,$$

where X, Y, $Z \in \Gamma(TM)$ and $\xi \in \Gamma(TM^\perp)$.

7.2.4 Proposition (Fundamental equations of an isometric immersion) *The first and second fundamental forms and the normal connection of an isometric immersion $f : (M,g) \to (\overline{M}, \overline{g})$ satisfy the following equations:*

$$\begin{aligned}
\langle \overline{R}(X,Y)Z, W\rangle &= \langle R(X,Y)Z, W\rangle \\
&\quad + \langle B(X,Z), B(Y,W)\rangle - \langle B(X,W), B(Y,Z)\rangle \quad \textit{(Gauss equation)} \\
(\overline{R}(X,Y)Z)^\perp &= (\nabla^\perp_X B)(Y,Z) - (\nabla^\perp_Y B)(X,Z) \quad \textit{(Codazzi-Mainardi equation)} \\
\langle \overline{R}(X,Y)\xi, \eta\rangle &= \langle R^\perp(X,Y)\xi, \eta\rangle - \langle [A_\xi, A_\eta]X, Y\rangle \quad \textit{(Ricci equation)}
\end{aligned}$$

where X, Y, Z, $W \in \Gamma(TM)$ and ξ, $\eta \in \Gamma(TM^\perp)$.

Proof. We first use the Gauss and Weingarten formulae (7.2.1), (7.2.3) to write

$$\begin{aligned}\overline{\nabla}_X\overline{\nabla}_Y Z &= \overline{\nabla}_X \nabla_Y Z + \overline{\nabla}_X B(Y,Z)\\ &= \nabla_X\nabla_Y Z + B(X,\nabla_Y Z) + \nabla^\perp_X(B(Y,Z)) - A_{B(Y,Z)}X.\end{aligned}$$

Then

$$\begin{aligned}\overline{R}(X,Y)Z &= \overline{\nabla}_X\overline{\nabla}_Y Z - \overline{\nabla}_Y\overline{\nabla}_X Z - \overline{\nabla}_{[X,Y]}Z\\ &= \nabla_X\nabla_Y Z + B(X,\nabla_Y Z) + \nabla^\perp_X(B(Y,Z)) - A_{B(Y,Z)}X\\ &\quad -\nabla_Y\nabla_X Z - B(Y,\nabla_X Z) - \nabla^\perp_Y(B(X,Z)) + A_{B(X,Z)}Y\\ &\quad -\nabla_{[X,Y]}Z - B(\nabla_X Y, Z) + B(\nabla_Y X, Z)\\ &= R(X,Y)Z + A_{B(X,Z)}Y - A_{B(Y,Z)}X\\ &\quad +\nabla^\perp_X B(Y,Z) - \nabla^\perp_Y B(X,Z).\end{aligned}$$

The tangential component of this equation is

$$(\overline{R}(X,Y)Z)^\top = R(X,Y)Z + A_{B(X,Z)}Y - A_{B(Y,Z)}X,$$

which is equivalent to the Gauss equation; the normal component is exactly the Codazzi-Mainardi equation.

Next, we use again the Gauss and Weingarten formulae to write

$$\begin{aligned}\overline{\nabla}_X\overline{\nabla}_Y \xi &= \overline{\nabla}_X\nabla^\perp_Y\xi - \overline{\nabla}_X A_\xi Y\\ &= \nabla^\perp_X\nabla^\perp_Y\xi - A_{\nabla^\perp_Y\xi}X - \nabla_X A_\xi Y - B(X, A_\xi Y).\end{aligned}$$

Then

$$\begin{aligned}\overline{R}(X,Y)\xi &= \overline{\nabla}_X\overline{\nabla}_Y\xi - \overline{\nabla}_Y\overline{\nabla}_X\xi - \overline{\nabla}_{[X,Y]}\xi\\ &= \nabla^\perp_X\nabla^\perp_Y\xi - A_{\nabla^\perp_Y\xi}X - \nabla_X A_\xi Y - B(X,A_\xi Y)\\ &\quad -\nabla^\perp_Y\nabla^\perp_X\xi + A_{\nabla^\perp_X\xi}Y + \nabla_Y A_\xi X - B(Y, A_\xi X)\\ &\quad -\nabla^\perp_{[X,Y]}\xi - A_\xi\nabla_X Y - A_\xi\nabla_Y X\\ &= R^\perp(X,Y)\xi + B(A_\xi X, Y) - B(X, A_\xi Y)\\ &\quad -(\nabla_X A_\xi)Y + (\nabla_Y A_\xi)X.\end{aligned}$$

It is easy to see that the tangential component of this equation yields again the Codazzi-Mainardi equation; we claim that the normal component is equivalent to the Ricci equation. In fact, it gives

$$\langle\overline{R}(X,Y)\xi,\eta\rangle = \langle R^\perp(X,Y)\xi,\eta\rangle + \langle B(A_\xi X, Y),\eta\rangle - \langle B(X, A_\xi Y),\eta\rangle,$$

where

$$\begin{aligned}\langle B(A_\xi X, Y),\eta\rangle - \langle B(X,A_\xi Y),\eta\rangle &= \langle A_\xi X, A_\eta Y\rangle - \langle A_\xi Y, A_\eta X\rangle\\ &= \langle A_\eta A_\xi X, Y\rangle - \langle A_\xi A_\eta X, Y\rangle\\ &= -\langle [A_\xi, A_\eta]X, Y\rangle.\end{aligned}$$

This completes the proof of the proposition. □

7.2.5 Corollary *If $(\overline{M},\overline{g})$ is a space form of curvature κ, then the fundamental equations for an isometric immersion $f:(M,g)\to(\overline{M},\overline{g})$ are given by:*

$$\begin{aligned}\langle R(X,Y)Z,W\rangle = -\kappa(\langle X,Z\rangle\langle Y,W\rangle - \langle Y,Z\rangle\langle X,W\rangle)\\ - \langle B(X,Z),B(Y,W)\rangle + \langle B(X,W),B(Y,Z)\rangle\end{aligned} \quad \textit{(Gauss equation)} \tag{7.2.6}$$

$$(\nabla^{\perp}_X B)(Y,Z) = (\nabla^{\perp}_Y B)(X,Z) \quad \textit{(Codazzi-Mainardi equation)} \tag{7.2.7}$$

$$\langle R^{\perp}(X,Y)\xi,\eta\rangle = \langle [A_\xi, A_\eta]X,Y\rangle, \quad \textit{(Ricci equation)} \tag{7.2.8}$$

where X, Y, Z, $W\in\Gamma(TM)$ and ξ, $\eta\in\Gamma(TM^{\perp})$.

Proof. The equations follow from the fact that the curvature tensor of $(\overline{M},\overline{g})$ is given by

$$\overline{R}(X,Y)Z = -\kappa(\langle X,Z\rangle Y - \langle Y,Z\rangle X),$$

see the end of section 4.2. □

7.2.9 Theorem (Fundamental theorem of submanifold geometry) *Let M be an n-dimensional manifold. Assume that we are given a Riemannian metric g on M, a rank k vector bundle E over M endowed with a Riemannian metric and a compatible connection ∇', and a symmetric E-valued tensor field B' on TM such that they satisfy the Gauss, Codazzi-Mainardi and Ricci equations for some real number κ. Then, for each point $p\in M$, there exists an open neighbourhood U of p in M and an isometric immersion f from U into the simply-connected space form of constant curvature κ and dimension $n+k$ such that g is the induced metric on U, $E|_U$ is isomorphic to the normal bundle of f, and B' and ∇' correspond respectively to the second fundamental form and the normal connection of f. Moreover, the isometric immersion f is locally uniquely defined up to congruence. If, in addition, M is assumed to be simply-connected, then the open set U can be taken to be all of M and f is uniquely defined up to congruence (however, in this case, f needs not be a global embedding).*

Proof. For simplicity, we prove the result for the case $\kappa = 0$ only. Define

$$A' : \Gamma(E)\to\Gamma(\mathrm{End}(TM)), \qquad \langle A'_\xi X,Y\rangle = \langle B'(X,Y),\xi\rangle$$

where X, $Y\in\Gamma(TM)$, $\xi\in\Gamma(E)$. Consider the rank $n+k$ Riemannian vector bundle $\bar{E} = TM\oplus E$, and define a connection $\bar{\nabla}$ on $\bar{E}$ as follows:

$$\bar{\nabla}_X Y = \nabla_X Y + B'(X,Y) \quad \text{and} \quad \bar{\nabla}_X\xi = -A'_\xi X + \nabla'_X\xi$$

for all X, $Y\in\Gamma(TM)$, $\xi\in\Gamma(E)$, where ∇ denotes the Levi-Cività connection of (M,g). It is easy to see that $\bar{\nabla}$ is compatible with the Riemannian metric on $\bar{E}$. One laboriously checks that the Gauss, Codazzi-Mainardi and Ricci compatibility equations precisely express the fact that $\bar{\nabla}$ is flat, namely, its curvature $\bar{R}$ vanishes everywhere. Therefore we can find a parallel orthonormal frame $\xi_1,\ldots,\xi_{n+k}$ of $\bar{E}$ defined on an open neighborhood U of p in M (compare exercise 6 of chapter 4). Consider the 1-forms $\theta_1,\ldots,\theta_{n+k}$ on M defined by $\theta_i(X) = \langle\xi_i, X\rangle$ for all $X\in\Gamma(TM)$, where the inner product is taken in $\bar{E}$. We compute for X, $Y\in\Gamma(TM)$:

$$\begin{aligned} d\theta_i(X,Y) &= X(\theta_i(Y)) - Y(\theta_i(X)) - \theta_i([X,Y])\\ &= \langle\bar{\nabla}_X\xi_i, Y\rangle + \langle\xi_i,\bar{\nabla}_X Y\rangle - \langle\bar{\nabla}_Y\xi_i, X\rangle - \langle\xi_i,\bar{\nabla}_X Y\rangle - \langle\xi_i,[X,Y]\rangle\\ &= \langle\xi_i,\nabla_X Y\rangle + \langle\xi_i, B'(X,Y)\rangle - \langle\xi_i,\nabla_X Y\rangle - \langle\xi_i, B'(Y,X)\rangle - \langle\xi_i,[X,Y]\rangle\\ &= 0,\end{aligned}$$

where we have used that each ξ_i is $\bar{\nabla}$-parallel, B' is symmetric and ∇ is torsionless. By shrinking U, if necessary, we can find smooth functions $f_1, \ldots, f_{n+k}$ on U such that $df_i = \theta_i$ for all i. We claim that $f = (f_1, \ldots, f_{n+k}) : U \to \mathbf{R}^{n+k}$ has the required properties.

For all $X \in \Gamma(TE)$, we have

$$\begin{aligned} \langle df(X), df(X) \rangle &= \sum_{i=1}^{n+k} df_i(X)^2 \\ &= \theta_i(X)^2 \\ &= \langle \xi_i, X \rangle^2 \\ &= g(X, X), \end{aligned}$$

since $\xi_1, \ldots, \xi_{n+k}$ is orthonormal, which shows that f is an isometric immersion. By shrinking U further, if necessary, we may assume that f is an embedding. Next, define a bundle isomorphism $F : \bar{E}|_U \to T\mathbf{R}^{n+k}|_{f(U)}$ by sending ξ_i to the ith element e_i of the canonical frame of $\mathbf{R}^{n+k}$. Note that

$$df(X) = \sum_{i=1}^{n+k} \langle \xi_i, X \rangle e_i = \sum_{i=1}^{n+k} \langle \xi_i, X \rangle F(\xi_i) = F\left(\sum_{i=1}^{n+k} \langle \xi_i, X \rangle \xi_i \right) = F(X)$$

for all $X \in \Gamma(TM)$, namely, F maps TM onto $df(TM)$. By construction, F maps a parallel orthonormal frame to a parallel orthonormal frame, so it preserves the metric and the connection. It follows that F maps E to the normal bundle νM and

$$F(\bar{\nabla}_X Y) = D_{F(X)} F(Y), \quad \text{and} \quad F(\bar{\nabla}_X \xi) = D_{F(X)} F(\xi)$$

for X, $Y \in \Gamma(TM)$ and $\xi \in \Gamma(E)$, where D denotes the Levi-Cività connection of $\mathbf{R}^{n+k}$; taking the tangent and normal components yields that F maps B' and ∇' respectively to the second fundamental form and normal connection of f. This finishes the proof of the local existence result.

Suppose next $f = (f_1, \ldots, f_{n+k}) : U \to \mathbf{R}^{n+k}$ is a given isometric immersion defined on a neighborhood U of p in M such that $E|_U$ is isomorphic to the normal bundle of f, and B' and ∇' correspond respectively to the second fundamental form and the normal connection of f. We first claim f is necessarily obtained from the above construction. Indeed, let $\bar{E} := f^* T\mathbf{R}^{n+k}$ be the vector bundle over U which is induced along f, namely, whose sections are exactly the vector fields aong f. Then $\{\xi_i := f^*(e_i)\}$ is a parallel orthonormal frame in $\bar{E}$ and the induced connection on $\bar{E}$ is flat. The assumptions on f and $E|_U$ imply that there is a bundle isomorphism $\bar{E} \cong TM|_U \oplus E|_U$ preserving metrics and connections. Finally, $df_i(X) = \langle df(X), e_i \rangle = \langle X, \xi_i \rangle$ for all i and all $X \in \Gamma(TM)$. Now for the uniqueness, note that if U is connected, the frame $\{\xi_i\}$ is uniquely determined up to an orthogonal transformation of $\bar{E}_p \cong \mathbf{R}^{n+k}$, and the functions f_i are uniquely determined up to an additive constant by the condition $df_i(\cdot) = \langle \xi_i, \cdot \rangle$, so f is uniquely determined up to a rigid motion of $\mathbf{R}^{n+k}$. Note that this result can be rephrased as saying that the immersion $f : U \to \mathbf{R}^{n+k}$ for U connected is uniquely specified by the initial values at p, that is $f(p) \in \mathbf{R}^{n+k}$ and $df_p \in \mathrm{Hom}(T_pM, \mathbf{R}^{n+k})$.

Finally, assume M is simply-connected. Given $q \in M$, connect p to q by a smooth curve γ and cover its image by finitely many connected open sets $U_1, \ldots, U_r$ such that there is an isometric immersion $U_i \to \mathbf{R}^{n+k}$ as above for each i, where $p \in U_1$, $q \in U_r$ and $U_i \cap U_{i+1}$ is non-empty and connected for $i = 1, \ldots, r-1$. Then f exists and is uniquely defined on a neighborhood of q and, in fact, on $\cup_{i=1}^r U_i$ by its initial values at p. Moreover, f is unchanged by a smooth homotopy of γ fixing the endpoints. Since M is assumed simply-connected, this shows that the value of f on q is independent of the choice of γ. Hence f is globally defined on M. □

7.3 The hypersurface case

Suppose that the codimension of M in $\overline{M}$ is one, and that both of these manifolds are oriented. Then a globally defined unit normal vector field ν can be defined on M; fix such ν. Then the second fundamental form can be viewed as real valued. Let $p \in M$. As the Weingarten operator $A_p = A_{\nu_p} : T_pM \to T_pM$ is self-adjoint, there exists a basis of T_pM consisting of eigenvectors of A_p with corresponding eigenvalues $\lambda_1(p), \ldots, \lambda_n(p)$, where $n = \dim M$. This defines functions $\lambda_1, \ldots, \lambda_n$ on M which are called the *principal curvatures* of M. The *multiplicity* of a principal curvature is the dimension of the corresponding eigenspace of A_p. The symmetric functions on the principal curvatures are invariants of the isometric immersion, up to sign in case of the symmetric functions of odd order (since the unit normal is unique up to sign only). Two significant instances of this invariants are the *mean curvature*

$$H := \frac{1}{n}\,\mathrm{trace}\,(A) = \frac{1}{n}(\lambda_1 + \cdots + \lambda_n),$$

and the *Gauss-Kronecker curvature*

$$K := \det(A) = \lambda_1 \cdots \lambda_n.$$

We specialize even more to the case in which $\overline{M} = \mathbf{R}^{n+1}$. Then the tangent spaces of $\mathbf{R}^{n+1}$ at its various points are canonically identified with $\mathbf{R}^{n+1}$ itself. The *Gauss map* of the immersion is the smooth map

$$g : M \to S^n,$$

where $g(p)$ is the unit vector $\nu_p \in S^n \subset \mathbf{R}^{n+1}$. Under the identifications, we can write

$$T_{\nu_p}S^n = (\mathbf{R}\nu_p)^\perp = T_pM.$$

It follows that the derivative of the Gauss map can be considered as a map $dg_p : T_pM \to T_pM$. Let $u \in T_pM$, and choose a smooth curve $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ and $\gamma'(0) = u$. Then

$$\begin{aligned} dg_p(u) &= \frac{d}{dt}\Big|_{t=0}(\nu \circ \gamma)(t) \\ &= \overline{\nabla}_u\nu \\ &= (\overline{\nabla}_u\nu)^\top \qquad (7.3.1) \\ &= -A_p(u), \end{aligned}$$

where we have used that $\langle \nu, \overline{\nabla}_u\nu\rangle = 0$ and Lemma 7.2.2.

Let u_i, u_j be eigenvectors of A_p of unit length associated to principal curvatures $\lambda_i(p)$, $\lambda_j(p)$, respectively; we may assume u_i and u_j are orthogonal (this is automatic if $\lambda_i(p) \neq \lambda_j(p)$). Then the Gauss equation (7.2.6) yields that the sectional curvature of the plane spanned by u_i, u_j is

$$-\langle R_p(u_i, u_j)u_i, u_j\rangle = \langle B_p(u_i, u_i), B_p(u_j, u_j)\rangle = \lambda_i\lambda_j. \qquad (7.3.2)$$

In the case $n = 2$, we recover (compare Proposition 4.6.1):

7.3.3 Theorem (Theorema Egregium of Gauss) *The Gaussian curvature of a surface in* $\mathbf{R}^3$ *is an intrinsic invariant; namely, it depends only on the first fundamental form (Riemannian metric).*

We close this section with an application that will be later generalied.

7.3.4 Theorem (Hadamard's convexity theorem) *Let $f : M \to \mathbf{R}^{n+1}$ be an immersion of a compact connected smooth manifold of dimension $n \geq 2$. Assume that the induced Riemannian metric on M has positive sectional curvature. Then M is diffeomorphic to S^n, f is an embedding, and $f(M)$ is a convex hypersurface (namely, the (smooth) boundary of a convex body) in $\mathbf{R}^{n+1}$.*

Proof. The positivity of the sectional curvature implies via the Gauss formula (7.3.2) that the product of any two principal curvatures has always the same sign. It follows that all principal curvatures have the same sign, namely, the second fundamental form is definite as a symmetric bilinear form, for any choice of normal vector at any point. This shows that we can continuously choose a unit normal vector field ξ along f such that A_ξ is definite positive at all points (in particular, M is already orientable). Owing to equation (7.3.1), this implies that the Gauss map $g : M \to S^n$ is a local diffeomorphism. Since M is compact and S^n is simply-connected, indeed g is a global diffeomorphism.

Next, for a fixed unit vector $v \in S^n \subset \mathbf{R}^{n+1}$ we consider the *height function* $h_v : M \to \mathbf{R}$ defined by $h_v(x) = \langle f(x), v\rangle$ for $x \in M$. It is clear that $(\operatorname{grad} h_v)_p$ is the orthogonal projection of v into $df(T_pM)$, so p is a critical point of h_v if and only if $v \in \nu_pM := df(T_pM)^\perp$ if and only if $v = \pm\xi_p$; this proves that h_v has exacly two critical points, as g is a diffeomorphism. Moreover, for a critical point p and X, $Y \in \Gamma(TM)$, we have (cf. exercise 13 of chapter 4):

$$\begin{aligned}
\operatorname{Hess}(h_v)(X,Y)_p &= X_p(Y(h_v)) \\
&= X_p\langle df(Y), v\rangle \\
&= \langle \nabla^f_{X_p} df(Y),\ v\rangle \\
&= \langle A_vX, Y\rangle_p.
\end{aligned}$$

Since A_v is definite, any critical point is isolated and a local maximum or local minimum. Since h_v must have a global maximum and a global minimum by compacteness of M, we deduce that for every $v \in S^n$ the height function h_v has exactly two critical points and

$$\min h_v \ \leq h_v(x) \leq \max\ h_v$$

where the first (resp. second) equality occurs if and only if x is the point of global minimum (resp. maximum). We deduce that f is injective and $f(M)$ is the boundary of a convex body. $\square$

7.4 Totally geodesic and totally umbilic submanifolds

A submanifold M of a Riemannian manifold $\bar{M}$ is called *totally geodesic* at a point $p \in M$ if the second fundamental form B vanishes at p, and it is called simply *totally geodesic* if B vanishes everywhere.

7.4.1 Proposition *For a submanifold M of $\bar{M}$, the following assertions are equivalent:*

a. *M is totally geodesic in $\bar{M}$;*
b. *every geodesic of M is a geodesic of $\bar{M}$;*
c. *the geodesic γ_v of $\bar{M}$ with initial velocity $v \in T_pM$ is contained in M for small time (and hence is a geodesic in M).*

Proof. Since B is symmetric, M is totally geodesic in $\bar{M}$ if and only $B(X,X) = 0$ for all $X \in \Gamma(TM)$ if and only if $B(v,v) = 0$ for all $v \in TM$. Gauss's formula (7.2.1) says that this is the case if and only if $\bar{\nabla}_X X = \nabla_X X$ for all $X \in \Gamma(TM)$. If this equation is true, plainly every geodesic in M will be a geodesic in $\bar{M}$. Conversely, assume every geodesic in M is a geodesic in

$\bar{M}$. Given $0 \neq v \in T_pM$, let γ_v be the geodesic in M with $\gamma_v'(0) = v$. Extend γ_v' to a smooth vector field $X \in \Gamma(TM)$ defined on a neighborhood of p. Since γ_v is also a geodesic of $\bar{M}$, we have $\bar{\nabla}_X X = 0 = \nabla_X X$. This proves the equivalence between (a) and (b). Next, if (b) holds, then the uniqueness of geodesics for given initial conditions says that all geodesics of $\bar{M}$ initially tangent to M come from geodesics of M, which implies (c). Finally, $\nabla_X X$ is the tangential component of $\bar{\nabla}_X X$ for $X \in \Gamma(TM)$, so a geodesic of $\bar{M}$ which is contained in M is a geodesic of M, which finishes the proof of the equivalence between (b) and (c). Note that the geodesic γ_v as in (c) is entirely contained in M, if M is complete. □

7.4.2 Corollary *A connected complete totally geodesic submanifold M of a Riemannian manifold $\bar{M}$ is completely characterized by T_pM for any given $p \in M$.*

Proof. In fact, it follows from the Hopf-Rinow theorem and Proposition 7.4.1 that $M = \overline{\exp}_p(T_pM)$, where $\overline{\exp}$ denotes the exponential map of $\bar{M}$. □

7.4.3 Proposition (Totally geodesic submanifolds of space forms) *The connected complete totally geodesic submanifolds of:*

a. $\mathbf{R}^n$ are the affine subspaces;

b. S^n are the great subspheres, namely, intersections of S^n with linear subspaces of $\mathbf{R}^{n+1}$;

c. $\mathbf{R}H^n$ are the intersections of hyperboloid model with linear subspaces of $\mathbf{R}^{1,n}$.

Proof. (a) Affine subspaces are clearly totally geodesic in $\mathbf{R}^n$. Since a totally geodesic submanifold is completely determined by its tangent space at a point, there can be no other examples. (b) Great circles of the subsphere are great circles of S^n, so this is a totally geodesic submanifold. The rest follows as in (a). The proof of (c) is similar. □

A submanifold M of a Riemannian manifold $\bar{M}$ is called *umbilic* in the direction of a normal vector ξ if the Weingarten operator A_ξ is a multiple of the identity operator, and it is called *totally umbilic* if every normal vector is umbilic; the latter property is equivalent to having

$$B(X,Y) = g(X,Y)\, H \tag{7.4.4}$$

for all X, $Y \in \Gamma(TM)$, where H is the mean curvature vector. This equation is equivalent to

$$\langle A_\xi X, Y\rangle = \langle X, Y\rangle\langle \xi, H\rangle$$

for all $\xi \in \Gamma(\nu M)$, X, $Y \in \Gamma(TM)$. Note that the minimal totally umbilic submanifolds are precisely the totally geodesic submanifolds. A totally umbilic submanifold with non-zero parallel mean curvature is called an *extrinsic sphere*.

7.4.5 Proposition *A totally umbilic submanifold of dimension at least two in a space form is an extrinsic sphere.*

Proof. Differentiate (7.4.4) with respect to $Z \in \Gamma(TM)$ and use $\nabla g = 0$ to get to get $(\nabla^\perp_Z B)(X,Y) = g(X,Y)\, \nabla^\perp_Z H$. Now the Codazzi equation 7.2.4 says that

$$g(X,Y)\, \nabla^\perp_Z H = g(Z,Y)\, \nabla^\perp_X H.$$

Since $\dim M \geq 2$, we can choose $Y \perp Z$ and $X = Y$ to deduce $\nabla^\perp_Z H = 0$. Since Z is arbitrary, H is parallel. □

7.4.6 Proposition (Totally umbilic submanifolds of space forms) *The connected complete non-totally geodesic totally umbilic submanifolds of dimension at least two in:*

a. $\mathbf{R}^n$ *are the round spheres;*

b. S^n *are the small subspheres, namely, intersections of* S^n *with non-linear affine subspaces of* $\mathbf{R}^{n+1}$*;*

c. $\mathbf{R}H^n$ *are the intersections of hyperboloid model with non-linear affine subspaces of* $\mathbf{R}^{1,n}$.

Proof. (a) Let $\iota : M \to \mathbf{R}^n$ be a connected non-totally geodesic totally umbilic submanifold of Euclidean space with $\dim M \geq 2$. Then $B = g\,H$ and $\nabla^\perp H = 0$. For $X \in \Gamma(TM)$, we compute

$$\bar\nabla_X \left(\iota + \frac{H}{||H||^2}\right) = X - \frac{1}{||H||^2} A_H X = 0.$$

Connectedness of M implies that it is contained in the hypersphere of radius $1/||H||$ and center $p + \frac{1}{||H||^2}H(p)$ for any $p \in M$. If M has codimension one and is complete, it must coincide with that hypersphere. If M has higher codimension, note that $A_\xi = 0$ for $\xi \perp H$. This implies that a parallel normal vector field which is orthogonal to H at one point must be constant. It follows that M is contained in the affine subspace containing p and parallel to the linear subspace spanned by T_pM and $H(p)$, for all $p \in M$. Now if M is complete then it coincides with the intersection of the above hypersphere with this affine subspace.

(b) Let $\iota : M \to S^n$ be a connected non-totally geodesic totally umbilic submanifold of the sphere with $\dim M \geq 2$. Let $\theta = \operatorname{arccot} ||H|| \in (0, \pi/2]$. For $X \in \Gamma(TM)$, denoting the Levi-Cività connection of S^n by $\bar\nabla$, we compute

$$\bar\nabla_X \left(\cos\theta\iota + \sin\theta\frac{H}{||H||}\right) = (\cos\theta - ||H||\sin\theta)X = 0.$$

If M is connected, this proves that M lies in the geodesic hypersphere of S^n of radius θ and center $n := \cos\theta p + \sin\theta\frac{H(p)}{||H||}$, for any $p \in M$. If M has codimension one and is complete, then it must coincide with this hypersphere, which is also the intersection of S^n with the affine hyperplane of $\mathbf{R}^{n+1}$ with normal n and distance $\cos\theta$ from the origin.

If M has higher codimension, let $\xi \perp H(p)$ and extend it to a parallel normal vector field $\hat\xi$ along M in S^n. Then, for all $X \in \Gamma(TM)$,

$$X(\hat\xi) = \bar\nabla_X\hat\xi + \langle X(\hat\xi), \iota\rangle\iota = -A_\xi X - \langle\hat\xi, X\rangle\iota = 0,$$

since $A_\xi = 0$, showing that $\hat\xi$ is constant in $\mathbf{R}^{n+1}$. This implies that M is contained in the hyperplane $\xi^\perp$. Another way to argue: note that for all $X \in \Gamma(TM)$,

$$X\left(\sin\theta\iota - \cos\theta\frac{H}{||H||}\right) = (\sin\theta + \cos\theta||H||)X \in \Gamma(TM)$$

and

$$X(Y) = \bar\nabla_X Y + \langle X(Y), \iota\rangle\iota = \nabla_X Y + \langle X, Y\rangle(H - \iota)$$

for $Y \in \Gamma(TM)$. This implies that the span of T_pM and $\sin\theta\, p - \cos\theta\,\frac{H(p)}{||H||} = \sin\theta(p - H(p))$ is a constant subspace E of $\mathbf{R}^{n+1}$ along M. Either way, we deduce from the completeness of M that it coincides with the intersection of S^n with the affine subspace $p + E$ of $\mathbf{R}^{n+1}$.

(c) Similar to case (b), but replacing trigonometric functions by their hyperbolic brothers and Euclidean space by Lorentzian space. □

Osculating spaces

Let M be a submanifold of a Riemannian manifold $\bar{M}$. The submanifold M is called *full* or *substantial* in $\bar{M}$ if M is not contained in a totally geodesic submanifold of $\bar{M}$ of dimension smaller than $\dim \bar{M}$; otherwise we say that the codimension can be reduced, or that there is a *reduction of the codimension* of M. The smallest number to which the codimension can be reduced is called the *substantial codimension* of M in $\bar{M}$.

In order to study the substantial codimension of submanifolds, we discuss a bit about osculating spaces. The *k-th osculating space* to M at $p \in M$, for $k = 1, 2, \ldots$, is the subspace $\mathcal{O}_p^k(M)$ of $T_p\bar{M}$ spanned by the first k derivatives at 0 of all smooth curves $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$. Here the higher derivatives of γ are defined by

$$\gamma'' = \frac{\bar{\nabla}}{dt}\gamma', \ \gamma''' = \frac{\bar{\nabla}}{dt}\gamma'', \text{ etc.}$$

Clearly, $\mathcal{O}_p^1(M) = T_pM$ and there is an increasing chain of subspaces

$$\mathcal{O}_p^1(M) \subset \mathcal{O}_p^2(M) \subset \cdots \subset T_p\bar{M} \tag{7.4.7}$$

for all $p \in M$. It follows from the Gauss equation that

$$\frac{\bar{\nabla}}{dt}\gamma' = \frac{\nabla}{dt}\gamma' + B(\gamma', \gamma')$$

where B denotes the second fundamental form of M in $\bar{M}$. Since B is symmetric, the subspace of ν_pM spanned by the image of B_p coincides with the subspace spanned by the image of B_p restricted to the diagonal of T_pM. We deduce that $\mathcal{O}_p^2(M)$ is spanned by $\mathcal{O}_p^1(M)$ and the image of B_p. Similarly, one sees that $\mathcal{O}_p^k(M)$ is spanned by all vectors of the form

$$X_1|_p, \ \bar{\nabla}_{X_1}X_2|_p, \ldots, \bar{\nabla}_{X_1} \cdots \bar{\nabla}_{X_{k-1}}X_k|_p$$

for $X_1, \ldots, X_k \in \Gamma(TM)$. The *k-th normal space* $\mathcal{N}_p^k(M)$ of M in $\bar{M}$ is the orthogonal complement of $\mathcal{O}_p^k(M)$ in $\mathcal{O}_p^{k+1}(M)$, so that

$$\mathcal{O}_p^k(M) \overset{\perp}{\oplus} \mathcal{N}_p^k(M) = \mathcal{O}_p^{k+1}(M).$$

Note that $v \perp \mathcal{N}_p^1(M)$ if and only if the Weingarten operator $A_v = 0$.

If $\dim \mathcal{O}_p^k(M)$ is independent of $p \in M$, then the collection of k-th osculating spaces to M at all points can be made into a vector subbundle $\mathcal{O}^k(M)$ of the vector bundle $T\bar{M}|_M$; if this is true for all k, then also the collection of k-th normal spaces to M at all points can be made into a vector subbundle $\mathcal{N}^k(M)$ of the normal bundle νM.

7.4.8 Lemma *For each $p \in M$, the chain (7.4.7) stabilizes at some $k_0 \geq 2$ (which may depend on p), namely,*

$$\mathcal{O}_p^{k_0-1}(M) \subsetneq \mathcal{O}_p^{k_0}(M) = \mathcal{O}_p^{k_0+1}(M) = \cdots$$

If k_0 does not depend on p, then $\mathcal{O}^{k_0}(M)$ is a parallel subbundle of $T\bar{M}|_M$ and $\mathcal{N}^1(M) \oplus \cdots \oplus \mathcal{N}^{k_0-1}(M)$ is a parallel subbundle of νM.

Proof. The chain stabilizes simply by dimensional reasons. An arbitrary smooth section ξ of $\mathcal{O}^k(M)$ is a sum of terms of the form $f\bar\nabla_{X_1}\cdots\bar\nabla_{X_{k-1}}X_k$, where $f\in C^\infty(M)$. Since

$$\bar\nabla_X(f\bar\nabla_{X_1}\cdots\bar\nabla_{X_{k-1}}X_k)=X(f)\bar\nabla_{X_1}\cdots\bar\nabla_{X_{k-1}}X_k+f\bar\nabla_X\bar\nabla_{X_1}\cdots\bar\nabla_{X_{k-1}}X_k,$$

we see that $\bar\nabla_X\xi\in\Gamma(\mathcal{O}^{k+1}(M))$ for all $X\in\Gamma(TM)$. Now if k_0 does not depend on p, then $\bar\nabla_X\Gamma(\mathcal{O}^{k_0}(M))\subset\Gamma(\mathcal{O}^{k_0}(M))$ for all $X\in\Gamma(TM)$, which is to say that $\mathcal{O}^{k_0}(M)$ is invariant under parallel transport in $T\bar M|_M$. This is equivalent to having that $\mathcal{N}^1(M)\oplus\cdots\oplus\mathcal{N}^{k_0-1}(M)$ is invariant under parallel transport in νM, which means that it is a parallel subbundle. □

In the remainder of this section, we assume that $\bar M$ is a space form.

7.4.9 Theorem (Erbacher) *Let M be an m-dimensional connected Riemannian submanifold of a space form $\bar M$. If $\mathcal{L}$ is a $\nabla^\perp$-parallel subbundle of νM containing $\mathcal{N}^1(M)$, and ℓ is the rank of $\mathcal{L}$, then there is a totally geodesic submanifold N of $\bar M$ of dimension $m+\ell$ that contains M.*

Proof. We consider separately the instances of space forms. The first case is $\bar M=\mathbf{R}^n$. Fix $p\in M$. It suffices to show that M is contained in the affine subspace $p+T_pM\oplus\mathcal{L}_p$ for some $p\in M$. Let γ be any piecewise smooth curve in $\bar M$ emanating from p and take any parallel normal vector field ξ along γ such that $\xi(0)\perp\mathcal{L}_p$. Since $\mathcal{L}$ is parallel along γ, we have that $\xi(t)\perp\mathcal{L}_{\gamma(t)}$ for all t. In particular, $\xi(t)\perp\mathcal{N}^1_{\gamma(t)}(M)$, so the Weingarten equation says that $\frac{\bar\nabla}{dt}\xi\equiv0$, namely, ξ is constant in $\mathbf{R}^n$ along γ. Since γ is arbitrary and M is connected, this means that M is contained in the orthogonal complement to the vector $\xi(0)$. Since $\xi(0)$ is an arbitrary vector in $\nu_pM\cap\mathcal{L}_p^\perp$, this case is done.

Consider next the case $\bar M=S^n(1)$; we reduce it to the previous case as follows. View M as a submanifold of $\mathbf{R}^{n+1}$ and consider the augmented vector bundle $\hat{\mathcal{L}}$ over M where $\hat{\mathcal{L}}_p=\mathcal{L}_p\oplus\mathbf{R}\,p$ for all $p\in M$. Note that $\hat{\mathcal{L}}$ is a subbundle of the normal bundle $\hat\nu M$ of M in $\mathbf{R}^{n+1}$ that contains the first normal bundle $\hat{\mathcal{N}}^1(M)$ of M in $\bar M$. Let $\hat\nabla^\perp$ be the normal connection on $\hat\nu M$. Both $\hat\nabla^\perp$ and $\nabla^\perp$ are induced by $\bar\nabla$, so they coincide on νM. Given that $\mathcal{L}$ is $\nabla^\perp$-parallel and the position vector $\mathbf{p}$ is $\hat\nabla^\perp$-parallel, we see that $\hat{\mathcal{L}}$ is $\hat\nabla$-parallel. By the previous case, M is contained in $(p+T_pM\oplus\hat{\mathcal{L}}_p)\cap S^n(1)$ for some $p\in M$, which is isometric to $S^{m+\ell}(1)$, as $p+\hat{\mathcal{L}}_p=\hat{\mathcal{L}}_p$ is a linear subspace of dimension $\ell+1$ of $\mathbf{R}^{n+1}$.

In case $\bar M=\mathbf{R}H^n$, by using the hyperboloid model the proof follows arguments similar to those in the previous two cases, where we use the canonical connection of Lorentz space $\mathbf{R}^{1,n}$. Indeed, as in the second case we view M as a submanifold of $\mathbf{R}^{1,n}$ and extend $\mathcal{L}$ to a subbundle $\hat{\mathcal{L}}$ of the normal bundle $\hat\nu M$ of M in $\mathbf{R}^{1,n}$ by adding the position vector field $\mathbf{p}$. We then prove, as in the first case, that M is contained in an affine subspace of $\mathbf{R}^{1,n}$ whose linear part, owing to $\langle\mathbf{p},\mathbf{p}\rangle=1$, is a Lorentz subspace isometric to $\mathbf{R}^{1,\ell}$

7.5 Focal points and the Morse index theorem

In this subsection, we state and prove the Morse index theorem for submanifolds of Riemannian manifolds. The discussion herein extends that in chapter 5 and, specially, Theorem 7.5.4 generalizes Theorem 5.5.3.

Let M be a submanifold of a Riemannian manifold $\bar M$. The restriction of the exponential map of $\bar M$ to the normal bundle of M is called the *normal exponential map* of M:

$$\exp^\perp:\nu M\to\bar M.$$

A critical value of $\exp^\perp$ is called a *focal point* of M. Note that this concept reduces to that of a conjugate point, in case M is a point.

7.5.1 Remark By the Sard-Brown theorem (see [Mil97]), the set of focal points of a submanifold has Lebesgue measure zero and hence is a nowhere dense set.

7.5.2 Proposition *A point $q \in \bar{M}$ is a focal point of a submanifold M if and only if there exists a geodesic $\gamma : [0,1] \to \bar{M}$ with $\gamma(0) = p \in M$, $\gamma(1) = q$, $\gamma'(0) = \xi \in \nu_p M$ and a Jacobi field J along γ such that $J(0) = u \in T_pM$ and $J'(0) + A_\xi u \in \nu_p M$.*

Proof. The point q is a critical value of $\exp^\perp$ if and only if it is in the image, namely, $q = \exp^\perp \xi$ for some $\xi \in \nu_p M$ and some $p \in M$, and the kernel of $d(\exp^\perp)_\xi$ is non-zero. Consider the geodesic $\gamma(s) = \exp_p(s\xi)$ for $s \in [0,1]$. Take a non-zero vector in $\ker d(\exp^\perp)_\xi$ represented by a smooth curve $\hat{\xi} : (-\epsilon, \epsilon) \to \nu M$, where $\hat{\xi}(0) = \xi$. This defines a smooth variation of γ through geodesics orthogonal to M:

$$H(s,t) = \exp^\perp_{c(t)}(s\hat{\xi}(t))$$

where $c(t) \in M$ is the footpoint of $\hat{\xi}(t)$ and $(s,t) \in [0,1] \times (-\epsilon, \epsilon)$. The associated variational vector field is a Jacobi field J along γ. Its initial conditions are (compare the proof of Proposition 5.4.4):

$$J(0) = \frac{\bar{\partial}}{\partial t}\Big|_{\substack{t=0\\s=0}} = c'(0) =: u,$$

and

$$\frac{\bar{\partial}}{\partial s}\Big|_{s=0} = d(\exp_{c(t)})_{0_{c(t)}}(\hat{\xi}(t)) = \hat{\xi}(t),$$

so

$$J'(0) = \overline{\nabla}_{\frac{\partial}{\partial s}} \frac{\bar{\partial}}{\partial t}\Big|_{\substack{t=0\\s=0}} = \overline{\nabla}_{\frac{\partial}{\partial t}} \frac{\bar{\partial}}{\partial s}\Big|_{\substack{t=0\\s=0}} = \bar{\nabla}_u \hat{\xi} = -A_\xi u + \nabla^\perp_u \hat{\xi},$$

completing the proof. □

A geodesic γ in a Riemannian manifold $\bar{M}$ which is perperndicular to a submanifold M at a point $p \in M$ is called an *M-geodesic*. A variational vector field along an M-geodesic which is associated to a variation through M-geodesics is called an *M-Jacobi field*. It follows from the proof of Proposition 7.5.2 that the space of M-Jacobi fields along an M-geodesic γ is the space of Jacobi fields J along γ that satisfy the initial conditions

$$J(0) \in T_pM \quad \text{and} \quad J'(0) + A_\xi J(0) \in \nu_p M,$$

where $p = \gamma(0)$ and $\xi = \gamma'(0)$. The *multiplicity* of a focal point $q = \gamma(s_0)$ to M along γ is the dimension of the kernel of $d(\exp^\perp)_{s_0\xi}$, which is also the dimension of the space of M-Jacobi fields along γ that vanish at s_0.

The Morse index theorem

Let M be a submanifold in a Riemannian manifold $\bar{M}$. Fix a unit speed M-geodesic $\gamma : [0,\ell] \to \bar{M}$ with $\gamma(0) = p \in M$. Denote by $\mathcal{V}$ the space of piecewise smooth vector fields Y along γ that satisfy the boundary conditions:

$$Y(0) \in T_pM, \quad Y'(0) + A_\xi Y(0) \in \nu_p M \quad \text{and} \quad Y(\ell) = 0,$$

where $\xi = \gamma'(0)$. Consider the index form I on $\mathcal{V}$ given by

$$I(X,Y) = -\langle A_\xi X, Y\rangle_0 + \int_0^\ell \langle X', Y'\rangle + \langle R(\gamma', X)\gamma', Y\rangle \, ds.$$

Note that $\frac{d^2}{dt^2}|_{t=0}E(\gamma_t) = I(Y,Y)$ for a variation of γ whose associated variational vector field Y lies in $\mathcal{V}$. It is not difficult to see that the kernel of I precisely consists of the M-Jacobi fields along γ that vanish at ℓ.

7.5.3 Lemma *Choose any subdivision* $0 = s_0 < s_1 < \cdots < s_n = \ell$ *such that* $\gamma|_{[s_{i-1},s_i]}$ *is minimizing for* $i = 1,\ldots,n$, $\gamma(s_1)$ *is not focal to* M *along* γ, *and* $\gamma(s_{i-1})$, $\gamma(s_i)$ *are not conjugate along* γ *for* $i = 2,\ldots,n$. *Then there is a* I*-orthogonal, vector space direct sum*

$$\mathcal{V} = \mathcal{V}^+ \oplus \mathcal{V}^-$$

where:

- $\mathcal{V}^+$ *is the subspace of* $\mathcal{V}$ *consisting of vector fields vanishing at* $s_0,\ldots,s_{n-1}$*;*
- $\mathcal{V}^-$ *is the subspace of* $\mathcal{V}$ *consisting of vector fields* X *that restrict to Jacobi fields along* $[s_{i-1},s_i]$ *for all* $i = 1,\ldots,n$.

Moreover, I *is positive definite on* $\mathcal{V}^+$. *It follows that the index (resp. nullity) of* I *on* $\mathcal{V}$ *is equal to the index (resp. nullity) of* I *on* $\mathcal{V}^-$*; in particular, it is finite.*

Proof. Let $X \in \mathcal{V}$. Since $\gamma(s_1)$ is not focal to M and $\gamma(s_{i-1})$ and $\gamma(s_i)$ are not conjugate points for $i \geq 2$ along γ, we can find $Y \in \mathcal{V}^-$ such that $Y(s_i) = X(s_i)$ for all $i = 0,\ldots,n$ (exercise 5 of chapter 5). Then $X - Y \in \mathcal{V}^+$. Clearly, $\mathcal{V}^+ \cap \mathcal{V}^- = \{0\}$.

Also, for $Y \in \mathcal{V}^-$ and $Z \in \mathcal{V}^+$, we have that Y is a Jacobi field (hence smooth) along $\gamma|_{[s_{i-1},s_i]}$ for $i = 1,\ldots,n$, so integration by parts allows us to rewrite the index form on $\mathcal{V}$ as (compare 5.4.1)

$$I(Y,Z) = -\langle A_\xi Y, Z\rangle_{s_0} - \sum_{i=1}^{n-1}\langle Y'(s_i^+) - Y'(s_i^-), Z\rangle + \int_0^\ell \langle -Y'' + R(\gamma',Y)\gamma', Z\rangle\, ds.$$

Since $Z(s_i) = 0$ for $i = 0,\ldots,n-1$, this formula shows that $I(Y,Z) = 0$.

Next we prove that I is positive definite on $\mathcal{V}^+$. Let $Z \in \mathcal{V}^+$. Since $\gamma|_{[s_{i-1},s_i]}$ is a minimizing geodesic and Z is the variational vector field associated to a variation that keeps $\gamma(s_{i-1})$ and $\gamma(s_i)$ fixed for $i = 1,\ldots,n$, we get that $I(Z,Z) \geq 0$. Suppose now, in addition, that $I(Z,Z) = 0$. For all $\tilde{Z} \in \mathcal{V}^+$ we have

$$0 \leq I(Z + \alpha\tilde{Z}, Z + \alpha\tilde{Z}) = 2\alpha I(Z,\tilde{Z}) + \alpha^2 I(\tilde{Z},\tilde{Z})$$

for all $\alpha \in \mathbf{R}$, which implies that $I(Z,\tilde{Z}) = 0$. Therefore Z is I-orthogonal to $\mathcal{V}^+$, and since it was already I-orthogonal to $\mathcal{V}^-$, we deduce that Z is a Jacobi field along γ. It follows that $Z = 0$.

The remaining assertions follow from the fact that $\mathcal{V}^-$ is finite-dimensional. □

7.5.4 Theorem (Morse) *Let* M *be a submanifold in a Riemannian manifold* $\bar{M}$. *Fix a unit speed* M*-geodesic* $\gamma : [0,\ell] \to \bar{M}$ *with* $\gamma(0) = p \in M$. *Then the index of* $I : \mathcal{V}\times\mathcal{V} \to \mathbf{R}$ *is finite and equals the sum of the multiplicities of focal points to* M *along* γ *of the form* $\gamma(s)$ *for some* $0 < s < \ell$.

7.5.5 Corollary *The set of focal points along an* M*-geodesic is discrete.*

Consider the restriction $\gamma_s := \gamma|_{[0,s]}$ for $s \in [0,\ell]$, the corresponding decomposition $\mathcal{V}_s = \mathcal{V}_s^+ \oplus \mathcal{V}_s^-$ as in Lemma 7.5.3, the associated index form I_s, and its index $\lambda(s)$. The proof of Theorem 7.5.4 is a consequence of Lemmata 7.5.6, 7.5.7, 7.5.8 and 7.5.10, in which we prove that λ is a left-continuous function with jumps precisely at the focal points.

7.5.6 Lemma $\lambda(s) = 0$ *for sufficiently small* $s > 0$.

Proof. This follows from the fact that γ_s is minimizing for sufficiently small $s > 0$. □

7.5.7 Lemma *$\lambda(s)$ is a non-decreasing function of $s \in [0, \ell]$.*

Proof. For $s' < s''$, we can view $\mathcal{V}_{s'} \subset \mathcal{V}_{s''}$ via the embedding

$$W \mapsto \overline{W}, \quad \text{where} \quad \overline{W}(s) = \begin{cases} W(s), & \text{if } 0 \le s \le s'; \\ 0, & \text{if } s' \le s \le s''; \end{cases}$$

and then $I_{s'}$ is the restriction of $I_{s''}$ to $\mathcal{V}_{s'}$, implying the result. □

7.5.8 Lemma *Given $s \in (0, \ell]$, we have $\lambda(s - \epsilon) = \lambda(s)$ for sufficiently small $\epsilon > 0$.*

Proof. Choose a subdivision of $[0, \ell]$ as in Lemma 7.5.3 such that $s \in (s_{i-1}, s_i]$ for some $i = 1, \ldots, n$. Since $\lambda(s) = 0$ for $s \in (0, s_1]$, we may assume $i \ge 2$. Since $\gamma(s_1)$ is not focal to M along γ, $d(\exp^{\perp})_{s_1\xi} : T_{s_1\xi}(\nu M) \to T_{\gamma(s_1)}\bar{M}$ is an isomorphism, where $\xi = \gamma'(0)$, so given $u \in T_{\gamma(s_1)}\bar{M}$, there exists a unique M-Jacobi field along γ_{s_1} whose value at s_1 is u. This, together with exercise 5 of chapter 5, shows that

$$\mathcal{V}_{s'}^{-} \cong T_{\gamma(s_1)}\bar{M} \oplus \cdots \oplus T_{\gamma(s_{i-1})}\bar{M} =: U \tag{7.5.9}$$

for all $s' \in (s_{i-1}, s]$. Note that $\lambda(s')$ is the index of $I_{s'}$ on U. Since $I_{s'}$, as a bilinear form on U, depends continuously on s', $I_{s'}$ is negative definite on any subspace of U on which I_s is negative definite, for sufficiently small $s - s' > 0$. This implies $\lambda(s') \ge \lambda(s)$ for sufficiently small $s - s' > 0$, and hence the desired result in view of Lemma 7.5.7. □

7.5.10 Lemma *Given $s \in (0, \ell)$, let $\nu(s)$ denote the nullity of I_s. Then $\lambda(s + \epsilon) = \lambda(s) + \nu(s)$ for sufficiently small $\epsilon > 0$.*

Proof. Choose a subdivision of $[0, \ell]$ as in Lemma 7.5.3 such that $s \in (s_{i-1}, s_i)$ for some $i = 1, \ldots, n$. Again, we may assume $i \ge 2$. Consider $I_{s'}$ as a bilinear form on U for $s' \in [s, s_i)$, where U is given as in (7.5.9). Note that I_s is positive definite on a subspace of dimension $\dim U - \lambda(s) - \nu(s)$. By continuity, also $I_{s'}$ is positive definite on that subspace for sufficiently small $s' - s > 0$. Therefore $\lambda(s + \epsilon) \le \lambda(s) + \nu(s)$ for sufficiently small $\epsilon > 0$.

To prove the reverse inequality, we start with linearly independent vector fields $X_1, \ldots, X_{\lambda(s)}$ in $\mathcal{V}_s$ spanning a subspace on which I_s is negative definite. Extend these vector fields over $\gamma_{s+\epsilon}$ by setting them equal to zero on $[s, s + \epsilon]$ as in Lemma 7.5.7. If $\epsilon > 0$ is sufficiently small, these extensions span a subspace of dimension $\lambda(s)$ of $\mathcal{V}_{s+\epsilon}$ on which $I_{s+\epsilon}$ is negative definite.

Next, by hypothesis we can find $\nu(s)$ linearly independent M-Jacobi fields $Y_1, \ldots, Y_{\nu(s)}$ along γ_s vanishing at s; extend them over $\gamma_{s+\epsilon}$ by zero. By making use of the technique of the theorem of Jacobi-Darboux 5.5.3, we can produce perturbations $\tilde{Y}_1, \ldots, \tilde{Y}_{\nu(s)} \in \mathcal{V}_{s+\epsilon}$ that span a subspace on which $I_{s+\epsilon}$ is negative definite. Since $X_1, \ldots, X_{\lambda(s)}, Y_1, \ldots, Y_{\nu(s)}$ were clearly linearly independent, we can also take the perturbations so that $X_1, \ldots, X_{\lambda(s)}, \tilde{Y}_1, \ldots, \tilde{Y}_{\nu(s)}$ are linearly independent. This completes the proof the lemma and of Theorem 7.5.4. □

7.6 Theory of isoparametric submanifolds

We say a Riemannian submanifold M of a space form $\bar{M}$ has *constant principal curvatures* if the principal curvatures along any locally defined parallel normal vector field are constant. If, in addition, M has flat normal bundle, then it is called an *isoparametric submanifold* of $\bar{M}$. Note that in case M has codimension one, both conditions reduce to simply requiring that M has constant principal curvatures. In view of the Fundamental Theorem 7.2.9, isoparametric submanifolds are sometimes said to be the submanifolds having the "simplest" local invariants.

Basic structure

Let M be an isoparametric submanifold of a space form $\bar{M}$. Since the normal bundle is flat ($R^\perp \equiv 0$), the Ricci equation (7.2.8) yields that, for every $p \in M$, $\{ A_\xi \mid \xi \in \nu_p M \}$ is a commutative family of symmetric endomorphisms of T_pM, hence simultaneously diagonalizable, say, with pairwise distinct eigenvalues $\lambda_1(\xi), \dots, \lambda_g(\xi)$ and common eigenspaces $E_1(p), \dots, E_g(p)$. Since the principal curvatures $\lambda_1(\xi), \dots, \lambda_g(\xi)$ are constant along any extension of ξ to a parallel normal vector field, we obtain g mutually orthogonal Frobenius distributions $E_1, \dots, E_g$ on M such that $TM = \bigoplus_{i=1}^g E_i$; each E_i is called a *curvature distribution*. The dimension of E_i is called a *multiplicity*. The restriction of λ_i to ν_pM is a linear functional, so there exists $v_i(p) \in \nu_pM$ such that $\lambda_i(\xi) = \langle \xi, v_i(p) \rangle$ for all $\xi \in \nu_pM$. This way we obtain g smooth normal vector fields $v_1, \dots, v_g$ along M, called *curvature normals*, which moreover are parallel, as:

$$\langle \nabla^\perp_X v_i, \xi \rangle = X\langle v_i, \xi \rangle - \langle v_i, \nabla^\perp_X \xi \rangle = X(\lambda_i(\xi)) = 0,$$

for every parallel normal vector field ξ and $X \in \Gamma(TM)$. Now for each $\xi \in \nu M$, the corresponding Weingarten operator satisfies

$$A_\xi|_{E_i} = \langle \xi, v_i \rangle \operatorname{id}_{E_i} \tag{7.6.1}$$

for $i = 1, \dots, g$; equivalently,

$$B(X_i, Y_j) = \langle X_i, Y_j \rangle \, v_i \tag{7.6.2}$$

for all $X_i \in E_i$, $Y_j \in E_j$ and i, $j = 1, \dots, g$. It follows from (7.6.1) that the case $g = 1$ precisely corresponds to the class of totally umbilic submanifolds of $\bar{M}$. Note that the substantial codimension of M equals the number of linearly independent curvature normals; this number is called the *rank* of M. We will always assume that M is full in $\bar{M}$, that is, not contained in a proper totally geodesic submanifold. It then follows that the curvature normals of M span the normal space at each point.

Another fundamental invariant of M is the covariant derivative of the second fundamental form. By taking derivatives and using the parallelism of the metric and the curvature normals, we obtain from (7.6.2) that

$$\nabla_{X_i} B(Y_j, Z_k) = \langle \nabla_{X_i} Y_j, Z_k \rangle \, (v_j - v_k) \tag{7.6.3}$$

for all $X_i \in E_i$, $Y_j \in E_j$, $Z_k \in E_k$ and i, j, $k = 1, \dots, g$. The Codazzi equation (7.2.7) is the symmetry of ∇B in all three arguments, which owing to (7.6.3), gives

$$\langle \nabla_{Z_k} X_i, Y_j \rangle \, (v_i - v_j) = \langle \nabla_{X_i} Y_j, Z_k \rangle \, (v_j - v_k) = \langle \nabla_{Y_j} Z_k, X_i \rangle \, (v_k - v_i).$$

Taking $i = j \neq k$ (in case $g \geq 2$) in the first equality shows that $\nabla_{X_i} Y_i \in \Gamma(E_i)$ for all X_i, $Y_i \in E_i$, namely, each curvature distribution E_i is *auto-parallel*; it follows that it is involutive and thus, by Frobenius theorem, integrable. Again, by auto-parallelism of E_i, its leaf through a point $p \in M$, denoted $S_i(p)$, is a totally geodesic submanifold of M. It follows from (7.6.1) that $S_i(p)$ is a totally umbilic submanifold of $\bar{M}$. In particular, in case $v_i = 0$ the distribution E_i is called the *nullity distribution* and its leaves are totally geodesic in $\bar{M}$.

We will be mostly concerned with the case of isoparametric submanifolds of Euclidean space. That this case contains the case of isoparametric submanifolds of spheres is the subject of the next proposition (see also Proposition 7.6.12).

7.6.4 Proposition *A Riemannian submanifold M of the unit sphere S^n of $\mathbf{R}^{n+1}$ is a rank k isoparametric submanifold of S^n if and only if it is a rank k isoparametric submanifold of $\mathbf{R}^{n+1}$.*

Proof. Let M be a submanifold of S^n. Every normal vector field ξ to M in S^n is also a normal vector field in $\mathbf{R}^{n+1}$. On the other hand, a normal vector field η to M in $\mathbf{R}^{n+1}$ can be written $\eta = \xi - f\mathbf{p}$, where ξ is the component of η tangential to S^n, $\mathbf{p}$ is the position vector, and f is a smooth function on M.

We will use the Weingarten formula. Denote by A, $\hat{A}$ the Weingarten operators of M, and by $\nabla^\perp$, $\hat{\nabla}^\perp$ the normal connections of M, viewed as a submanifold of S^n, $\mathbf{R}^{n+1}$, respectively. Denote by ∇ the Levi-Cività connection of S^n and take $X \in \Gamma(TM)$. Note that

$$\nabla_X \xi = X(\xi) - \langle X(\xi), \mathbf{p}\rangle \mathbf{p} = X(\xi),$$

since $\langle X(\xi), \mathbf{p}\rangle = -\langle \xi, X(\mathbf{p})\rangle = -\langle \xi, X\rangle = 0$. It follows that

$$\begin{aligned} -\hat{A}_\eta X + \hat{\nabla}^\perp_X \eta &= X(\eta) \\ &= X(\xi) - X(f)\mathbf{p} - fX \\ &= \nabla_X \xi - X(f)\mathbf{p} - fX \\ &= -A_\xi X + \nabla^\perp_X \xi - X(f)\mathbf{p} - fX. \end{aligned}$$

Comparing tangent and normal components, we obtain that

$$\hat{\nabla}^\perp_X \eta = \nabla^\perp_X \xi - X(f)\mathbf{p} \quad \text{and} \quad \hat{A}_\eta X = A_\xi X + fX.$$

We deduce from these equations that η is $\hat{\nabla}^\perp$-parallel if and only if ξ is $\nabla^\perp$-parallel and f is constant; and, in this case, the eigenvalues of $\hat{A}_\eta$ are of the form $\lambda_i(\xi) + \mu$, where $\lambda_i(\xi)$ is an eigenvalue of A_ξ and μ is a constant, with the same eigenspaces. Note that the curvature normals v_i, $\hat{v}_i$ of M as an isoparametric submanifold of S^n, $\mathbf{R}^{n+1}$, resp., are related by $\hat{v}_i = v_i - \mathbf{p}$. □

Parallel foliation

Let M be an isoparametric submanifold of a space form $\bar{M}$. For a fixed parallel normal vector field ξ, a fundamental construction is the *parallel map*

$$\pi_\xi : M \to \bar{M}, \qquad \pi_\xi(x) = \exp^\perp \xi(x),$$

namely, the restriction of the normal exponential map along ξ. For simplicity, in the sequel we assume $\bar{M} = \mathbf{R}^{n+k}$, where $n = \dim M$.

Now $\pi_\xi(x) = x + \xi(x)$ for $x \in M$. Using the canonical parallelism of $\mathbf{R}^{n+k}$, the differential of this map is $\mathrm{id} - A_\xi$, so its kernel is $\bigoplus\{ E_i \mid \langle \xi, v_i\rangle = 1\}$. Since π_ξ has constant rank, its image M_ξ is a submanifold of $\bar{M}$ of dimension $n - \dim\ker d(\pi_\xi)_p$ for $p \in M$. The map $\pi_\xi : M \to M_\xi$ is a submersion, and M_ξ is called a *parallel manifold* in case $\dim M_\xi = n$, or a *focal manifold* in case $\dim M_\xi < n$. We thus see that the *focal set* of M, namely the subset of $\bar{M}$ consisting of all focal points of M along normal geodesics, decomposes into focal manifolds, and M_ξ is a focal manifold precisely if $\ker d\pi_\xi$ is non-zero, in which case π_ξ is called a *focal map*.

Since $\pi := \pi_\xi : M \to M_\xi$ is a submersion, there is an orthogonal decomposition $TM = \mathcal{H} \oplus \mathcal{V}$, where $\mathcal{V}_p = \ker d\pi_p$ and $d\pi_p : \mathcal{H}_p M \to T_{\pi(p)}(M_\xi)$ is an isomorphism. Since $d\pi_p = \mathrm{id} - A_\xi$, we can view $T_pM = T_{\pi(p)}(M_\xi)$ and then $\nu_{\pi(p)}(M_\xi) = \nu_p M \oplus \mathcal{V}_p$.

7.6.5 Lemma *The connected components of the level sets of $\pi : M \to M_\xi$ are totally geodesic submanifolds of M. Moreover, the Weingarten operators of M preserve the decomposition $TM = \mathcal{H} \oplus \mathcal{V}$.*

Proof. The first assertion is to be a consequence of the Codazzi equation (7.2.7); cf. third equality below. Let $U, V \in \Gamma(\mathcal{V})$, $X \in \Gamma(\mathcal{H})$. Since $\nabla^\perp \xi = 0$, $A_\xi U = U$, $A_\xi V = V$, and using that A_ξ is a self-adjoint operator, we have:

$$\begin{aligned} \langle \nabla_U V, X \rangle &= \langle \nabla_U (A_\xi V), X \rangle \\ &= \langle (\nabla_U A_\xi) V + A_\xi (\nabla_U V), X \rangle \\ &= \langle (\nabla_X A_\xi) V, U \rangle + \langle \nabla_U V, A_\xi X \rangle \\ &= \langle \nabla_X (A_\xi V) - A_\xi (\nabla_X V), U \rangle + \langle \nabla_U V, A_\xi X \rangle \\ &= \langle \nabla_U V, A_\xi X \rangle, \end{aligned}$$

proving that $\langle \nabla_U V, d\pi(X) \rangle = 0$ and hence that $\mathcal{V}$ is auto-parallel.

Further, since ξ is parallel, the Ricci equation (7.2.8) says that the Weingarten operators of M at p commute with A_{ξ_p}. Therefore they commute with $d\pi_p$ and thus preserve its kernel $\mathcal{V}_p$, for all $p \in M$. Since they are symmetric endomorphisms of T_pM, they also preserve $\mathcal{H}_p$. □

7.6.6 Remark In fact, it is not hard to show that any component of the level set $\pi^{-1}(\pi(p))$ in Lemma 7.6.5 is an isoparametric submanifold of $\nu_{\hat{p}} M_\xi$, where $\hat{p} = p + \xi(p)$ (cf. Exercise 15). Using as a tool the normal holonomy of focal manifolds, one can work harder and see that those level sets are connected and indeed *homogeneous* isoparametric submanifolds [BCO16, § 4.3.3]. These are called *slices* of the given isoparametric submanifold.

Any smooth curve in M_ξ admits a locally defined lifting to a horizontal smooth curve in M (cf. exercise 19 of chaper 3). Let $\gamma : [a, b] \to M$ be a horizontal smooth curve and put $\hat{\gamma} = \pi \circ \gamma$. Since $\nu_{\gamma(t)} M \subset \nu_{\hat{\gamma}(t)} M_\xi$ for all t, any normal vector field η to M along γ can be also considered as a normal vector field $\hat{\eta}$ to M_ξ aong $\hat{\gamma}$.

7.6.7 Proposition (Tube formula) *For $v \in \nu_p M \subset \nu_{\pi(p)} M_\xi$, let $\hat{A}_v$ denote the Weingarten operator of M_ξ. Then*

$$\hat{A}_v = A_v \circ ((\mathrm{id}_{T_pM} - A_{\xi_p})|_{\mathcal{H}_p})^{-1}.$$

Proof. Using the canonical parallelism of $\mathbf{R}^{n+k}$ and the Weingarten formula (7.2.3), we can write

$$-\hat{A}_{\hat{\eta}(t)} \hat{\gamma}'(t) + \nabla^\perp_{\hat{\gamma}'(t)} \hat{\eta} = \hat{\eta}'(t) = \eta'(t) = -A_{\eta(t)} \gamma'(t) + \nabla^\perp_{\gamma'(t)} \eta,$$

where $\hat{\nabla}^\perp$ denotes the normal connection of M_ξ. Since γ is horizontal, we know from Lemma 7.6.5 that $A_{\eta(t)} \gamma'(t) \in \mathcal{H}_{\gamma(t)} = T_{\hat{\gamma}(t)}(M_\xi)$. It follows that $A_{\eta(t)} \gamma'(t) = \hat{A}_{\hat{\eta}(t)} \hat{\gamma}'(t)$. Now we need only remark that $\hat{\gamma}'(t) = d\pi_{\gamma(t)}(\gamma'(t)) = \gamma'(t) - A_{\xi(\gamma(t))} \gamma'(t)$. □

7.6.8 Corollary *The submanifold M_ξ has constant principal curvatures. Moreover, if M_ξ is a parallel manifold ($\dim M_\xi = \dim M$), then it is isoparametric.*

Proof. It follows from the tube formula that the principal curvatures of M_ξ at $v \in \nu_{\pi(p)} M_\xi$ are of the form

$$\frac{\lambda_i(v)}{1 - \lambda_i(\xi_p)}$$

for $i = 1, \ldots, g$, where the λ_i are the principal curvatures of M. Therefore M_ξ has constant principal curvatures.

It also follows from the proof of Proposition 7.6.7 that $\nabla^\perp_{\hat\gamma'(t)}\hat\eta = \nabla^\perp_{\gamma'(t)}\eta$. In case $\dim M_\xi = \dim M$, we have $\nu_{\pi(p)}(M_\xi) = \nu_p M$ implying that $\hat R^\perp = R^\perp = 0$. □

For $p \in M$, denote by $F_M(p)$ the subset of $\bar M$ consisting of focal points of M relative to p. Note that $F_M(p)$ consists of the points $q \in p + \nu_p M$ such that $d\pi_{q-p} = \mathrm{id} - A_{q-p}$ has not full rank.

7.6.9 Corollary *If M_ξ is a parallel manifold, then $p + \nu_p M = \hat p + \nu_{\hat p} M_\xi$ and $F_M(p) = F_{M_\xi}(\hat p)$ for all $p \in M$, where $\hat p = p + \xi(p)$.*

Proof. The first assertion is a consequence of the facts that $\nu_p M = \nu_{\hat p} M_\xi$ and $\hat p \in p + \nu_p M$. For the second one, the tube formula (7.6.7) yields that

$$\begin{aligned} \mathrm{id} - \hat A_{q-\hat p} &= (\mathrm{id} - A_{\xi(p)} - A_{q-\hat p})(\mathrm{id} - A_{\xi(p)})^{-1} \\ &= (\mathrm{id} - A_{q-p})(\mathrm{id} - A_{\xi(p)})^{-1}, \end{aligned}$$

for all $q \in p + \nu_p M$, whence we see that $\mathrm{id} - \hat A_{q-\hat p}$ is invertible if and only if so is $\mathrm{id} - A_{q-p}$, as desired. □

The Coxeter group

We next describe the Coxeter group associated to a complete isoparametric submanifold M of Euclidean space.

The affine normal space $p + \nu_p M$ meets the focal set $F_M(p)$ along the union of the affine hyperplanes $H_i(p) = p + \{\, \xi \in \nu_p M \mid \langle \xi, v_i(p)\rangle = 1 \,\}$ corresponding to non-zero curvature normals, called *focal hyperplanes* with respect to p. For each focal hyperplane $H_i(p)$, the orthogonal reflection of $p + \nu_p M$ on the hyperplane $H_i(p)$ will be denoted by $\tilde r_i^p$. We will show that the group generated by all the $\tilde r_i^p$ is a finite Coxeter group.

Recall that, in general, an (abstract) *Coxeter group* is a finitely presented group

$$\langle r_1, \ldots, r_n \mid (r_i r_j)^{m_{ij}} = 1 \rangle$$

where $m_{ii} = 1$ and $m_{ij} \geq 2$ if $i \neq j$, and the condition $m_{ij} = \infty$ means that no relation of the form $(r_i r_j)^m$ is imposed. The number n is called the *rank* of the Coxeter group. In 1934, H. S. M. Coxeter proved that every finite group generated by orthogonal reflections on hyperplanes in an Euclidean space is a Coxeter group, whereas in 1935 he proved that every finite Coxeter group admits a faithful representation as group generated by reflections on an Euclidean space and classified the finite Coxeter groups. They fall into: three families of increasing rank A_n, B_n, D_n; one family of rank two, $I_2(p)$; and six exceptional groups, E_6, E_7, E_8, F_4, H_3 and H_4.

Note that for each non-zero curvature normal v_i of M, the leaf $S_i(p)$ of E_i passing through p is the hypersphere in $p + E_i(p) + \mathbf{R}\, v_i(p)$ of center $c_i(p) = p + \frac{v_i(p)}{||v_i||^2}$ and radius $1/||v_i||$. Let $a_i : M \to M$ denote the map that restricts to the antipodal map of $S_i(x)$ for all $x \in M$. Namely, a_i is given by the parallel map π_{η_i} where $\eta_i = 2\frac{v_i}{||v_i||^2}$. This shows that a_i is an involutive diffeomorphism of M (but not an isometry).

From the fact that the normal bundle is globally flat■1■, for p, $q \in M$ we have a well-defined parallel transport map $\tau_{p.q} : \nu_p M \to \nu_q M$. Let $\tilde\tau_{p,q} : p + \nu_p M \to q + \nu_q M$ the associated affine parallel transport map. The parallelism of the curvature normals in fact implies that

$$\tilde\tau_{p,q}(F_M(p)) = F_M(q). \tag{7.6.10}$$

■1■!

7.6.11 Lemma *For all $p \in M$ and $i = 1, \ldots, g$, it holds that*

$$\tilde{r}_i^p(p) = a_i(p) \quad \text{and} \quad \tilde{r}_i^p = \tilde{\tau}_{p,a_i(p)}.$$

Proof. The first assertion follows from the fact that $v_i(p)$ points in the radial direction of $S_i(p)$. To prove the second one, first consider the parallel transport from p to $a_i(p)$ in the normal bundle to $S_i(p)$ in $p + \mathbf{R}\, v_i(p) \oplus E_i(p)$; this map clearly takes $v_i(p)$ to $v_i(a_i(p)) = -v_i(p)$. Since $p+\mathbf{R}\, v_i(p) \oplus E_i(p)$ is totally geodesic in $\bar{M}$ and $S_i(p)$ is invariant under the Weingarten operators of M, it follows from the Weingarten formula that the above map is the restriction of $\tilde{\tau}_{p,a_i(p)}$. Finally, if ξ is a normal vector field to M along a curve γ in $S_i(p)$, which is parallel and everywhere normal to $v_i(p)$, then, due to (7.6.1),

$$\frac{\bar{\nabla}\xi}{dt} = -A_{\xi(t)}\gamma'(t) + \frac{\bar{\nabla}^{\perp}\xi}{dt} = -\langle \xi(t), v_i(\gamma(t))\rangle\, \gamma'(t) = 0,$$

namely, ξ is constant in $\bar{M}$. This shows that $\tilde{\tau}_{p,a_i(p)}$ is the identity on $v_i(p)^{\perp}$, and finishes the proof. □

In view of Corollary 7.6.9, equation (7.6.10), $a_i(p) = \pi_{\eta_i}(p)$, and $M = M_{\eta_i}$, we now have that

$$\begin{aligned}
\tilde{\tau}_{p,a_i(p)}(F_M(p)) &= F_M(a_i(p)) \\
&= F_{M_{\eta_i}}(\pi_{\eta_i}(p)) \\
&= F_M(p)
\end{aligned}$$

for all $p \in M$. Due to Lemma 7.6.11, this says that $\tilde{r}_i^p$ acts on $p + \nu_p M$ by permuting the focal hyperplanes $H_i(p)$. Since there are only finitely many focal hyperplanes, this implies that the group W^p generated by all the $\tilde{r}_i^p$ is finite. Owing to the above quoted result of Coxeter, we deduce that W^p is a Coxeter group, called the *Coxeter group of M at p*. Note that the dependence on the point $p \in M$ is not very important, since $\tilde{\tau}_{p,q}$ conjugates W^p to W^q, so the conjugation class is uniquely defined and denoted simply by W. It is also usual to see W as a Coxeter group acting on the linear space $\nu_p M$. Note that the rank of W as a Coxeter group is the same as the rank of M as an isoparametric submanifold.

Decomposition theorems

Let M be a connected complete isoparametric submanifold of Euclidean space. The following remark is very important for the results in this subsection. Since the Coxeter group W associated to M is a finite group of orthogonal transformations of $p+\nu_p M$, it must have a fixed point (namely, the center of mass of any orbit). This means that there is a non-zero vector in $\bigcap_i H_i(p)$, so a non-zero parallel normal vector field ζ such that $\langle \zeta, v_i\rangle = 0$ for all non-zero curvature normals v_i.

7.6.12 Proposition *A connected complete isoparametric submanifold M of $\bar{M} = \mathbf{R}^n$ admits a splitting $M = N \times E_0$ such that $N = M \cap E_0^{\perp}$ is an isoparametric submanifold (of the same rank as M) of a sphere of dimension $n - \dim E_0$, and E_0 is the nullity distribution of M. Moreover, M and N have the same Weyl group.*

Proof. We denote the zero curvature normal by v_0, if it is present. Let ζ be a parallel normal vector field such that $\langle \zeta, v_i\rangle = 1$ for all $i \neq 0$, as above. The differential of the parallel map π_ζ has kernel equal to $\mathcal{D} = \bigoplus_{i\neq 0} E_i$. By Lemma 7.6.5, this distribution is auto-parallel. Since

$TM = \mathcal{D} \oplus E_0$ is an orthogonal decomposition, we have $\nabla_X U \in \Gamma(E_0)$ for $X \in \Gamma(\mathcal{D})$ and $U \in \Gamma(E_0)$. As a curvature distribution, E_0 is auto-parallel, so the Gauss formula yields

$$\bar{\nabla}_X U = \nabla_X U \in \Gamma(E_0)$$

for $X \in \Gamma(TM)$ and $U \in \Gamma(E_0)$, which implies that the distribution E_0 is constant along M as a subspace of Euclidean space; since M is complete, the leaf of the distribution through $p \in M$ is the affine subspace $p + E_0$.

Since M is ruled by affine subspaces parallel to the constant subspace E_0, it immediately follows that $M = N \times E_0$ where $N = M \cap E_0^\perp$ and N is connected. It remains to be seen that N is isoparametric in $E_0^\perp$. Note that $TN = \mathcal{D}|_N$ and $\nu N = \nu M|_N$. Since $E_0^\perp$ is totally geodesic in $\bar{M}$, we see from the Weingarten formula that the normal connection of M in $\bar{M}$ restricts to the normal connection of N in $E_0^\perp$, and the Weingarten operators of M (leave N invariant and) restrict to the Weingarten operators of N. We deduce that N is isoparametric in $E_0^\perp$. Since $d\pi_\eta(TN) = 0$ and N is connected, the map π_ζ is a constant c, which gives that N is contained in the sphere of center c and radius $||\zeta||$ in $E_0^\perp$.

The last assertion is true because the nullity distribution does not contribute to the Weyl group. □

7.6.13 Corollary *For a connected complete isoparametric submanifold M of $\mathbf{R}^n$, the following are equivalent:*

a. All curvature normals are non-zero.
b. M is contained in a round sphere of $\mathbf{R}^n$.
c. M is compact.

Proof. In the notation of the proposition: if all curvature normals are non-zero, then $M = N$ is contained in a sphere; complete isoparametric submanifolds of Euclidean space are always closed, so they are compact if contained in a sphere; by the proposition, M can be compact only if E_0 is trivial. □

Let $\bar{M}_1$ and $\bar{M}_2$ be Riemannian manifolds, and let M_i be a submanifold of $\bar{M}_i$ for $i = 1, 2$. The *extrinsic product* of M_1 and M_2 is the product $M_1 \times M_2$ viewed as a submanifold of the Riemannian product $\bar{M}_1 \times \bar{M}_2$.

An isoparametric submanifold M of Euclidean space $\mathbf{R}^n$ is said to be *reducible* if M is the extrinsic product of isoparametric submanifolds $M_i \subset \mathbf{R}^{n_i}$ for $i = 1, 2$ ($n = n_1 + n_2$), where M_1, M_2 are not points; note that in this case the Coxeter group of M is the product of the Coxeter groups of M_1 and M_2. Otherwise, we say that M is *irreducible.*

Let W denote a Coxeter group faithfully represented as a group generated by reflections acting on an Euclidean space V. The group W is called *reducible* if there exists a non-trivial decomposition $V = V_1 \oplus V_2$ into W-invariant subspaces. Note that in this case W is isomorphic to a product $W_1 \times W_2$ where W_i is a Coxeter group acting on V_i, for $i = 1, 2$.

7.6.14 Proposition *Let M be a compact isoparametric submanifold of $\mathbf{R}^n$ with Coxeter group W. Then M is reducible if and only if W is reducible.*

Proof. Assume W is reducible, namely, $W = W_1 \times W_2$ where W_i acts on $\mathbf{R}^{n_i}$ and $n = n_1 + n_2$. We want to prove that M is reducible. By applying a translation, we may assume that M passes through the origin of $\mathbf{R}^n$. Owing to Corollary 7.6.13, we know that all curvature normals of M are non-zero.

The set of generators of W splits as a union of the set of generators of W_1 and those of W_2; there is a corresponding splitting of the set of curvature normals into two sets V_1 and V_2. Note that these two sets V_1 and V_2 of vectors are mutually orthogonal. Let ζ be a parallel normal vector field to M in $\mathbf{R}^n$ such that $\langle \zeta, v_i \rangle = 1$ for every curvature normal v_i. Decompose $\zeta = \zeta_1 + \zeta_2$ where ζ_i lies in the span of V_i. Then, for each $i = 1, 2$, the number $\langle \zeta_i, v_j \rangle$ equals 1 or 0 according to whether v_j lies in V_i or not.

Define the distributions $\mathcal{D}_i = \ker(\mathrm{id} - A_{\zeta_i})$ and $\nu_i M = \sum_{v_j \in \mathsf{V}_i} \mathbf{R} v_j$, and put $V_i = \mathcal{D}_i \oplus \nu_i M$ for $i = 1, 2$. Note that there is a g- and B-orthogonal decomposition $TM = \mathcal{D}_1 \oplus \mathcal{D}_2$, where each $\mathcal{D}_i$ is parallel in M and invariant under its Weingarten operators. We claim that V_1 is constant as a subspace of $\mathbf{R}^n$ along M. Let $X \in \Gamma(TM)$, $Y \in \Gamma(\mathcal{D}_1)$, $\xi \in \Gamma(\nu_1 M)$. We easily compute that

$$\bar{\nabla}_X (Y + \xi) = \nabla_X Y + B(X, Y) - A_\xi X + \nabla^\perp_X \xi$$

lies in $\Gamma(V_1)$, proving the claim. Similarly, V_2 is constant as a subspace of $\mathbf{R}^n$ along M. Let $\mathbf{R}^{n_i}$ be the linear subspace of $\mathbf{R}^n$ given by V_i at the origin $0 \in M$ for $i = 1, 2$. Then $\mathbf{R}^n = \mathbf{R}^{n_1} \oplus \mathbf{R}^{n_2} = \mathbf{R}^{n_1} \times \mathbf{R}^{n_2}$, and we put $M_i = M \cap \mathbf{R}^{n_i}$ for $i = 1, 2$; note that M_i is the integral manifold of $\mathcal{D}_i$ through the origin. Similarly to the above computation, we easily see that $\mathcal{D}_1$ is constant as a subspace of $\mathbf{R}^n$ along $M \cap (\{p_1\} \times \mathbf{R}^{n_2})$ for $p_1 \in M_1$, and $\mathcal{D}_2$ is constant as a subspace of $\mathbf{R}^n$ along $M \cap (\mathbf{R}^{n_1} \times \{p_2\})$ for $p_2 \in M_2$. It follows that the integral manifolds of $\mathcal{D}_1$ (resp. $\mathcal{D}_2$) are all of the form $M_1 + p_2 = M \cap V_1$ (resp. $p_1 + M_2 = M \cap V_2$), which gives that $M = M_1 \times M_2$.

As in the proof of Proposition 7.6.12, one sees that M_i is (compact) isoparametric in $\mathbf{R}^{n_i}$. The submanifolds M_1 and M_2 are integral manifolds of the auto-parallel distributions $\mathcal{D}_1$ and $\mathcal{D}_2$, therefore they are totally geodesic submanifolds of M. Note that the normal bundle of M_i in V_i is the restriction of $\nu_i M$ to M_i, so the Coxeter group of M_i is indeed W_i. □

We gather from Propositions 7.6.12 and 7.6.14 that every complete isoparametric submanifold of Euclidean space splits an extrinsic product of its Euclidean factor and a number of irreducible compact isoparametric submanifolds (with irreducible Coxeter groups).

7.7 Examples and classification of isoparametric submanifolds

Isoparametric hypersurfaces

From Proposition 7.6.12 we recover Levi-Cività [LC37] and B. Segre's[Seg38] result that the number g of principal curvatures of an isoparametric hypersurface of Euclidean space is at most two, and it is either a hyperplane, a hypersphere or the boundary of a tube of constant radius around an affine subspace. Cartan extended Segre's bound on g to hyperbolic spaces and obtained a similar classification. In fact, Cartan studied isoparametric hypersurfaces systematically in a remarkable series of four papers [Car38, Car39a, Car39b, Car40] during the years 1938-40, and pointed out that isoparametric hypersurfaces in spheres are much more interesting and difficult objects of study.

From Proposition 7.6.14 we see that an isoparametric hypersurface in $S^n(1)$ with $g = 2$ must be a product $S^{m_1}(r_1) \times S^{m_2}(r_2)$, where $r_1^2 + r_2^2 = 1$; the family $\{S^{m_1}(\cos t) \times S^{m_2}(\sin t)\}_{t \in [0, \pi/2]}$ comprises an isoparametric family in the unit sphere S^n, where $n = m_1 + m_2 + 1$. The principal curvatures are easily seen to be

$$\lambda_1 = \cot t, \ \lambda_2 = -\tan t = \cot(t + \frac{\pi}{2}),$$

with (arbitrary) multiplicities m_1 and m_2. In particular, the focal hypersurfaces are points, corresponding $t = 0$ and $t = \pi/2$, and $t = \pi/4$ is the only parameter value corresponding to a minimal hypersurface; this hypersurface was found by W. K. Clifford in 1873 and is today known as the *Clifford torus*. Note that this foliation is the orbital foliation obtained from the standard action of $\mathbf{SO}(m_1+1)\times\mathbf{SO}(m_2+1)$ on $\mathbf{R}^{m_1+1}\times\mathbf{R}^{m_2+1}$.

Cartan constructed four examples of isoparametric foliations in spheres with $g=3$, all with uniform multiplicity equal to 1, 2, 4 or 8, and then proved that there are no other examples with $g=3$. Those hypersurfaces are all homogeneous. The simplest example lives in the unit sphere in the 5-dimensional Euclidean vector space V of traceless real symmetric 3×3 matrices equipped with the inner product $\langle X,Y\rangle=\mathrm{tr}(XY)$, which we describe as follows. There is an action of the Lie group $G=\mathbf{SO}(3)$ on V given by conjugation, namely, $g\cdot X=gXg^{-1}$ for $g\in G$ and $X\in V$. The isoparametric foliation of the unit sphere S^4 of V consists of conjugation classes of matrices of norm 1. The conjugation class of $X\in S^4$ is a compact submanifold; indeed it is the image of the immersion

$$g\in\mathbf{SO}(3)\mapsto gXg^{-1}\in S^4$$

which becomes injective after factoring $\mathbf{SO}(3)$ by the centralizer $Z_G(X)$ of X, which is a closed subgroup. Each symmetric matrix is conjugate to a diagonal matrix, so we can parametrize such classes by diagonal matrices. The centralizer of a diagonal (resp. arbitrary) matrix is discrete if and only if the matrix has pairwise distinct entries (resp. eigenvalues). Consider the following orthonormal basis of V:

$$e_1=\frac{1}{\sqrt6}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix},\quad e_2=\frac{1}{\sqrt2}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix},$$

$$e_3=\frac{1}{\sqrt2}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix},\quad e_4=\frac{1}{\sqrt2}\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix},\quad e_5=\frac{1}{\sqrt2}\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}.$$

Then the diagonal matrices in S^4 can be parametrized by the geodesic

$$\gamma(t)=\cos t e_1+\sin t e_2.$$

The matrix $\gamma(t)$ has distinct eigenvalues if and only if $t\neq k\pi/3$ for $k\in\mathbf{Z}$; for such a value of t, the conjugation class (orbit) M_t is 3-dimensional, and the tangent space $T_{\gamma(t)}M_t$ is spanned by e_3, e_4, e_5. In fact, denote by E_{ij} the matrix with coefficient 1 at position (i,j) and 0 elsewhere, and put $X_{ij}=E_{ij}-E_{ji}$. The one-parameter subgroup

$$g_s=\begin{pmatrix}\cos s&\sin s&0\\-\sin s&\cos s&0\\0&0&1\end{pmatrix}$$

of $\mathbf{SO}(3)$ yields the following tangent vector at $p=\gamma(t)$:

$$\begin{aligned}\frac{d}{ds}\Big|_{s=0}g_spg_s^{-1}&=X_{12}p-pX_{12}\\&=[X_{12},\gamma(t)]\\&=\frac{1}{\sqrt6}\cos t[X_{12},E_{11}+E_{22}-2E_{33}]+\frac{1}{\sqrt2}\sin t[X_{12},E_{11}-E_{22}]\\&=(-2\sin t)e_3;\end{aligned}$$

we call this vector $X_{12}p$. Similarly,

$$X_{23}p=(-2\sin t)e_4,\ X_{13}p=(-2\sin t)e_5.$$

Now

$$\xi = -\gamma'(t) = \sin t e_1 - \cos t e_2$$

is a unit normal vector to M_t in S^4. We extend ξ to a normal vector field along $s \mapsto g_s p$ by putting $\xi(s) = dg_s(\xi) = g_s\xi$ and then

$$\begin{aligned} A_\xi(e_3) &= \frac{-1}{2\sin t} A_\xi(X_{12}p) \\ &= \frac{1}{2\sin t} (\bar{\nabla}_{X_{12}p}\xi)^\perp \\ &= \frac{1}{2\sin t} \left(\frac{d}{ds}\Big|_{s=0} g_s \xi g_s^{-1} \right)^\top \\ &= \frac{1}{2\sin t} (+\sin t[X_{12}, e_1] - \cos t[X_{12}, e_2]) \\ &= \cot t\, e_3. \end{aligned}$$

Similarly,

$$A_\xi(e_4) = \cot(t + \frac{\pi}{3})e_4 \quad \text{and} \quad A_\xi(e_5) = \cot(t + \frac{2\pi}{3})e_5.$$

Therefore the principal curvatures are

$$\lambda_1 = \cot t,\ \lambda_2 = \cot(t + \frac{\pi}{3}),\ \lambda_3 = \cot(t + \frac{2\pi}{3}),$$

with corresponding curvature distributions spanned by e_3, e_4 and e_5, respectively. Note that $M_{\pi/6}$ is a minimal hypersurface of S^4, called the *Cartan hypersurface*. Any conjugation class meets $\gamma(t)$ for some $t \in [0, \pi/3]$, since we can always permute the eigenvalues of a diagonal matrix by conjugating it by a suitable orthogonal matrix (called a *permutation matrix*!). The interior points $\gamma(t)$ for $t \in (0, \pi/3)$ have pairwise distinct eigenvalues and hence discrete centralizers, namely, the group of diagonal matrices with ±1 entries. The endpoints $\gamma(0)$ and $\gamma(\pi/3)$ are matrices with an eigenvalue of multiplicity two, so its centralizers are larger, namely, the block subgroups $\mathbf{S}(\mathbf{O}(2)\mathbf{O}(1))$ and $\mathbf{S}(\mathbf{O}(1)\mathbf{O}(2))$ of $\mathbf{SO}(3)$, respectively. The focal manifolds $M_+ = M_0$ and $M_- = M_{\pi/3}$ are antipodal Veronese surfaces diffeomorphic to $\mathbf{R}P^2$. In particular, the multiplicities of the isoparametric family $\{M_t\}_{t\in[0,\pi/3]}$ are $m_1 = m_2 = 1$.

There is a beautiful, unified way to generalize the above example to include all examples with $g = 3$ discovered by Cartan. The standard embeddings of the projective spaces $\mathbf{F}P^n$, where $\mathbf{F}$ is one of the four normed division algebras over $\mathbf{R}$, namely, $\mathbf{R}$, $\mathbf{C}$, $\mathbf{H}$ (quaternions) and $\mathbf{Ca}$ (Cayley numbers; here n must be 2), are constructed as follows. Let V be the space $\mathrm{Herm}_\epsilon(n, \mathbf{F})$ be the space of $n \times n$ Hermitian matrices with coefficients in $\mathbf{F}$ and constant trace ϵ; ϵ is usually taken to be equal to 0 or 1. A one-dimensional subspace of $\mathbf{F}^{n+1}$ is identified with the orthogonal projection onto it, namely, an idempotent element in V; this realizes $\mathbf{F}P^n$ as the real algebraic smooth variety $M_+ = \{\, x \in V \mid x^2 = x \,\}$. Note that $\dim \mathbf{F}P^n = dn$ and $\dim V = (n-1)(dn+2)/2$, where $d = 1$, 2, 4 or 8, according to $\mathbf{F} = \mathbf{R}$, $\mathbf{C}$, $\mathbf{H}$ or $\mathbf{Ca}$. The squared Euclidean norm in V is $||x||^2 = \mathrm{trace}(x^2)$, so M_+ is contained in the unit sphere $S(V)$ of V. It can be shown that the tubes of constant radius $r \in [0, \pi/3]$ around M_+ in $S(V)$ comprise an isoparametric foliation of $S(V)$, where the tube with $r = \pi/3$ corresponds to the antipodal embedding $M_- = -M^+$ of $\mathbf{F}P^n$, and $M_\pm$ are the focal manifolds. These foliations are respectively homogeneous under the compact Lie groups $G = \mathbf{SO}(n)$, $\mathbf{SU}(n)$, $\mathbf{Sp}(n)$ and $\mathbf{F}_4$. The representations of the group G on V are given in the first three cases by $\rho(g)x = gxg^*$, where $x \in V$, $g \in G$ and x^* denotes the transpose conjugate matrix of x, and in the fourth case by the 26-dimensional representation of $\mathbf{F}_4$.

Later Cartan discusses the case $g = 4$ and shows there are only two examples where the multiplicities of principal curvatures are all equal, namely, one in S^5 and one in S^9. Towards the end of his third paper on the subject, Cartan asks three questions, one of which asking whether every isoparametric hypersurface in a sphere is homogeneous. It is clear that any orbit of codimension one in S^n of a closed subgroup of $\mathbf{SO}(n+1)$ has constant principal curvatures and is thus isoparametric. Hsiang and Lawson [HL71] classified connected closed subgroups of $\mathbf{SO}(n+1)$ whose principal orbits have codimension one in S^n. It turns out that the actions of such groups which are maximal connected, in the sense that they are not contained in a larger connected group with the same orbits, precisely coincide with the isotropy representations of symmetric spaces of rank two. Takagi and Takahashi [TT72] refer to [HL71] and note that it implies a classification of homogeneous isoparametric hypersurfaces in spheres. They relate the geometric invariants to the invariants of the corresponding symmetric spaces and list their multiplicities. In particular, they find examples with $g = 4$ and different multiplicities, for instance the orbits of the isotropy representation of the oriented Grassmann manifold of two-planes in $\mathbf{R}^{n+3}$ is an isoparametric submanifold of S^{2n+1} with $g = 4$ and multiplicities $m_1 = 1$, $m_2 = n - 1$.

After Cartan, the subject of isoparametric hypersurfaces in spheres remained dormant until the work of Takagi and Takahashi, and the short note of Nomizu [Nom73], in which he proved that the focal manifolds of an isoparametric family are always minimal submanifolds. Around the same time, Münzner did very influential work, published in the two papers [Mue80, Mue81] much later in 1981-2. In the first paper, he proved that there are exactly two focal manifolds. In the second paper, using delicate topological arguments based on the fact that a compact isoparametric submanifold of a sphere decomposes the sphere into a union of two closed ball bundles over the focal manifolds, Münzner proved the striking result that the only possible values of g are 1, 2, 3, 4 and 6, namely, the same values obtained from the homogeneous examples.

In 1975, Ozeki and Takeuchi [OT75] surprised the community of researchers in the field by exhibiting examples of inhomogeneous isoparametric hyperusfaces in spheres. These examples were later systematized and generalized by Ferus, Karcher and Münzner [FKM81], who associated examples with $g = 4$ to representations of Clifford algebras, most of which are inhomogeneous.

The classification problem of isoparametric hypersurfaces in spheres starts with the determination of the possible multiplicities (m_1, m_2). Cartan had already solved the problem for $g \leq 3$. In case $g = 6$, Abresch [Abr83] proved that only $(1, 1)$ and $(2, 2)$ are possible; note that indeed there are homogeneous examples with those multiplicities. The case $g = 4$ was the most involved and, after the efforts of many mathematicians, it was finally completed by Stolz [Sto99] who, in a topological tour de force, proved that the possibilities are exactly those that appear either in the homogeneous examples or in the Clifford examples of Ferus, Karcher and Münzner.

Isoparametric hypersurfaces with $g = 6$ and $(m_1, m_2) = (1, 1)$ must indeed be homogeneous by the work of Dorfmeister and Neher [DN85]. Their proof depends on an intricate algebraic calculation, and it seems very difficult to extend their approach to the case $(m_1, m_2) = (2, 2)$. More recently, the work of Cecil, Chi and Jensen [CCJ07], Immervoll [Imm08] and Chi [Chi12] shows that isoparametric hypersurfaces with $g = 4$ must be either homogeneous or one of the known inhomogeneous examples, with the possible exception of $(m_1, m_2) = (7, 8)$.

There have been attempts to simplify Dorfmeister-Neher's result and to extend it to the case $(g, m_1, m_2) = (6, 2, 2)$ [Miy09, Miy13, Miy15, Sif16].

General structure of isoparametric hypersurfaces of spheres

Let M be a compact isoparametric hypersurface of S^{n+1}. For $p \in M$ and a unit normal vector $\xi \in \nu_p M$, consider the normal geodesic $\gamma(t) = \cos t\, p + \sin t\, \xi$ for $t \in [0, 2\pi]$. Then γ meets the parallel and focal manifolds orthogonally. Since the codimension of M in $\mathbf{R}^{n+2}$ is two, in this case the Coxeter group is a dihedral group $\mathbf{D}_g$ (with $2g$ elements) with $g \geq 3$ in case M is irreducible, or $\mathbf{Z}_2 \times \mathbf{Z}_2 = \mathbf{D}_2$ or $\mathbf{Z}_2 = \mathbf{D}_1$ otherwise. It follows that the multiplicities satisfy the periodicity condition $m_i = m_{i+2}$ (indices modulo g); in particular, M has uniform multiplicities if g is odd. It also implies that the focal distances are equidistributed along the image of γ, and hence the principal curvatures can be written (cf. Exercise 12)

$$\cot d,\ \cot\left(d + \frac{\pi}{g}\right), \ldots, \cot\left(d + (g-1)\frac{\pi}{g}\right).$$

Isoparametric submanifolds

Palais and Terng [PT87] extended Takagi and Takahashi's remark to state that the principal orbits of the isotropy representation of a symmetric space are compact isoparametric submanifolds, see chapter **??**. In the same paper, using the classification of Dadok [Dad85] they also characterized the compact homogeneous isoparametric submanifolds of Euclidean space as being exactly those orbits. There remained the inhomogeneous isoparametric submanifolds to be understood. In 1991, invoking the theory of Tits buildings, Thorbergsson proved the deep result that every compact connected full irreducible isoparametric submanifold of Euclidean space with codimension at least 3 is homogeneous, showing thus that the FKM-examples are the only inhomogeneous ones, always in codimension 2. Thorbergsson's theorem has been reproved by Olmos [Olm93] using canonical connections and normal holonomy, and by Heintze and Liu [HL99]; the latter proof in fact also applies to the infinite dimensional case, cf. add. notes.

Marked Coxeter graph

Let M be a connected compact full isoparametric submanifold of an Euclidean sphere. It follows from equation (7.6.2) that the focal hyperplanes in $p + \nu_p M$ together with the multiplicities $m_i = \dim E_i$ for $i = 1, \ldots, g$, determine the second fundamental form, as an abstract symmetric bilinear form, up to passing to a parallel submanifold. In turn, the focal hyperplanes are already determined by the Weyl group, up to scaling of the ambient metric. Thus the Weyl group together with the multiplicities essentially determine the second fundamental form; such data is usually encoded in the form of a Coxeter graph with multiplicities, as follows.

Let W be the Coxeter group of M acting on $p + \nu_p M$ for some $p \in M$. A connected component of the complement of the union of the focal hyperplanes in $p + \nu_p M$ is called a *Weyl chamber*. The *Coxeter graph* of W is constructed by fixing a Weyl chamber $\mathcal{C}$ and taking as vertices the *walls* of $\mathcal{C}$, i.e. hyperplanes bounding $\mathcal{C}$. Note that these correspond to the generators $r_1, \ldots, r_n$ of W. Associated to each wall is a curvature distribution and the corresponding multiplicity, which we write on top of the vertex; this is the marking. Since the multiplciities are preserved under the action of W, the marking already determines all multiplicities. The two vertices corresponding to generators r_i and r_j are linked by an edge if and only if the corresponding walls are not perpendicular, in which case we write the number m_{ij} on top of the edge (recall that $(r_i r_j)^{m_{ij}} = 1$ is a relation in W; for simplicity, in case $m_{ij} = 3$ one usually writes nothing and the number 3 remains implicit). It turns out that W is irreducible if and only if its Coxeter graph is connected; in this case $\mathcal{C}$ is a simplicial cone and its Coxeter graph has $n = \dim \nu_p$ vertices. The isomorphism type of the Coxeter graph is independent of the chosen Weyl chamber, as W acts simply transitively on the set of Weyl chambers, and determines W up to isomorphism. The Coxeter graph together with the marking is called the *marked Coxeter graph* of the isoparametric submanifold.

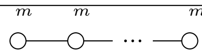
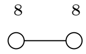
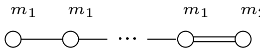
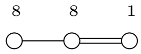
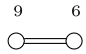
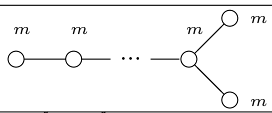
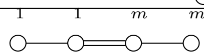
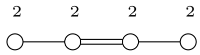
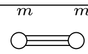
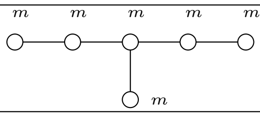
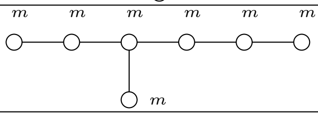
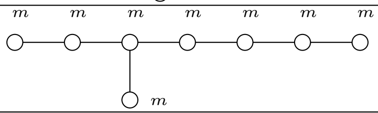
Type	Diagram	Multiplicities
$A_n,\ n \geq 2$		$m = 1, 2, 4$
A_2		–
$(BC)_n,\ n \geq 2$		m_1 m_2 1 k 2 $2,\ 2k+1$ 4 $1,\ 5,\ 4k+3$
$(BC)_3$		–
$(BC)_2$		–
$D_n,\ n \geq 4$		$m = 1, 2$
F_4		$m = 1, 2, 4, 8$
		–
G_2		$m = 1, 2$
E_6		$m = 1, 2$
E_7		$m = 1, 2$
E_8		$m = 1, 2$

Table 7.7.1: Coxeter graphs of homogeneous isoparametric submanifolds.

Recall that the connected compact full isoparametric hypersurfaces of Euclidean spaces are exactly the round hyperspheres of arbitrary radius, which have Coxeter graph of type A_1. Münzner's result quoted above says that the number g of principal curvatures of M is 1, 2, 3, 4 or 6. It follows that a rank 2 compact isoparametric submanifold of Euclidean space has Coxeter group of type A_1, $A_1 \times A_1$, A_2, B_2 or G_2. Due to Remark 7.6.6, for a compact isoparametric submanifold M of rank $n \geq 3$, any subgraph of the Coxeter graph of M which is obtained by removing some vertices of the graph of M and all edges linking to those vertices is the Coxeter graph of some slice of M. This fact shows that the admissible Coxeter graphs (groups) of isoparametric submanifolds if Euclidean space are A_n $(n \geq 1)$, B_n $(n \geq 2)$, D_n $(n \geq 4)$, E_n $(n = 6,\ 7,\ 8)$, F_4 and G_2. These are called *crystallographic Coxeter groups*, since they leave invariant a lattice. They are also called *Weyl groups*, since they appear in the theory of compact semisimple Lie algebras. In Table 7.7.1 we also list the possible multiplicities but only in the homogeneous case.

The following rigidity result shows that a homogeneous isoparametric submanifold M is completely characterized by the values of the second fundamental form B and its covariant derivative ∇B at one point $p \in M$. It is almost true that M is already determined by B, for the only exception are the adjoint orbits of the compact Lie groups $\mathbf{Spin}(2n+1)$ and $\mathbf{Sp}(n)$, whose marked

Coxeter graphs are isomorphic (with uniform multiplicity 2). For this reason, the theorem is more interesting in the infinite dimensional case, where it is also valid [GH12].

7.7.1 Theorem (Gorodski-Heintze) *Let M and M' be two connected complete full homogeneous isoparametric submanifolds of Euclidean spaces V and V', respectively. Assume there is an isometry $f : V \to V'$ and points $p \in M$, $p' \in M'$ such that $f(p) = p'$, $df_p(T_pM) = T_{p'}M'$, $df_p(B_p(u,v)) = B'_{p'}(df_p(u), df_p(v))$ and $df_p(\nabla_u B(v,w)) = \nabla_{df_p(u)} B'_{p'}(df_p(v), df_p(w))$, for all u, v, $w \in T_pM$, where B and B' denote the second fundamental forms of M and M', respectively. Then $f(M) = M'$.*

7.8 Additional notes

§1 In complex analysis of one variable, Liouville's theorem says that a bounded entire function is constant. Bernstein (1915-17) proved an analogous result in differential geometry, namely, if the graph of a function $f : \mathbf{R}^2 \to \mathbf{R}$ of class $\mathcal{C}^2$ is a minimal surface in $\mathbf{R}^3$, then the graph is a plane. He then posed the classical Bernstein problem, namely, whether the same result also holds for real functions of $n > 2$ variables. In terms of differential equations:

> **(Classical) Bernstein problem**: *Let the function $f : \mathbf{R}^n \to \mathbf{R}$ of class $\mathcal{C}^2$ be a solution of*
> $$\sum_{i=1}^n \frac{\partial}{\partial x_i}\left(\frac{\partial f/\partial x_i}{\sqrt{1+||\mathrm{grad} f||^2}}\right) = 0.$$
> *Must f be a linear function?*

Part of the importance of the Bernstein problem is that it has a direct bearing on the existence of minimal cones and singularities of minimal hypersurfaces in $\mathbf{R}^{n+1}$. The answer to the problem was proved to be affirmative in the cases $n = 3$ by de Giorgi (1965), $n = 4$ by Almgren (1966), and $n \leq 7$ by Simons (1968), and apparently there was some hope to extend the result to all dimensions. However, in 1969 Bombieri, de Giorgi and Giusti [EBG69] constructed a counter-example for $n = 8$, which yields a counter-example in each dimension $n > 8$ by a standard construction, closing the problem. The complete solution of the Bernstein problem turned out to involve a good deal of geometric measure theory and non-linear analysis.

§2 Let M be an isoparametric submanifold of $\bar{M} = \mathbf{R}^{n+k}$. Using the Coxeter group associated to M in an essential way, Terng proved in [Ter85] that M is the level set of a so called *isoparametric map* $F : \mathbf{R}^{n+k} \to \mathbf{R}^k$, namely, a map $F = (F_1, \ldots, F_k)$ admitting regular values and such that:

(i) the Laplacians ΔF_i are constant along the level sets of F, for $i = 1, \ldots, k$;

(ii) The inner products $\langle \mathrm{grad} F_i, \mathrm{grad} F_j \rangle$ are constant along the level sets of F, for all i, $j = 1, \ldots, k$;

(iii) The Lie brackets $[\mathrm{grad} F_i, \mathrm{grad} F_j]$ are linear combinations with constant coefficients of
$$\mathrm{grad} F_1, \ldots, \mathrm{grad} F_k$$
along the level sets of F, for all i, $j = 1, \ldots, k$.

(In case $k = 1$, conditions (i) and (ii) were classically referred to as expressing the constancy of the differential parameters ΔF_1 and $||\mathrm{grad} F_1||^2$ of F_1 along its level sets, hence the name *isoparametric*. Condition (iii) is a kind of integrability and is void in case $k = 1$.) Moreover, Terng showed that F can be taken polynomial. It follows that every connected isoparametric submanifold of Euclidean space is an open subset of a complete properly embedded isoparametric submanifold, which in addition is a real algebraic submanifold of Euclidean space. It is easy to check that, conversely, the regular level sets of an arbitrary isoparametric map are isoparametric submanifolds. The regular levels of F are exactly the parallel manifolds of M, and the singular levels are the focal manifolds of M. The resulting partition of $\bar{M}$ is called an *isoparametric foliation*, and it provides an important example of *singular Riemannian foliation* [Ale04].[2].

§3 The theory of isoparametric submanifolds of Euclidean space was extended to separable Hilbert spaces by Terng in [Ter89]. The local differential geometry of submanifolds in Euclidean spaces generalizes without much effort to Hilbert space. One is thus tempted to use the same definition, namely, constancy of principal curvatures along parallel normal vector fields and flat normal bundle. This works if one restricts to the category of *proper Fredholm* submanifolds of Hilbert space, that is, those submanifolds of Hilbert space whose normal exponential map is a proper Fredholm map. In practice, this says that such submanifolds have finite codimension and compact (self-adjoint) Weingarten operators. Terng generalized the whole structure theory of isoparametric submanifolds to Hilbert space, including the Coxeter group, which is now an (infinite) affine Weyl group. The structure is now more involved also for the reason that the distribution of nullity does not have to split off. On one hand, there is a remarkable family of examples of isoparametric foliations of Hilbert space coming from isotropy representations of (infinite-dimensional) affine Kac-Moody symmetric spaces. On the other hand, examples of FKM-type can be also be constructed in Hilbert space (without resorting to polynomials!, though [TT95]). Thorbergsson's theorem was extended to Hilbert space by Heintze and Liu, who proved that a connected complete full irreducible isoparametric submanifold of Hilbert space of rank at least 2 is extrinsically homogeneous [HL99]; however, little is known about the group acting transitively on that submanifold. The classification problem, even in the homogeneous case, is wide open, for there is no standard theory of infinite-dimensional Lie groups and their affine representations that one can apply. A recent contribution is [GH12], which characterizes such manifolds by the values of the second fundamental form B and its covariant derivative ∇B at one point (cf. Theorem 7.7.1), and proposes a strategy to the classification, namely, first obtain restrictions on ∇B (those on B are already known) and then compare with the known examples.

7.9 Exercises

[3]

1 Let V be an inner product space. For a basis $(v_1, \ldots, v_n)$ of V, let A be the matrix of a linear transformation $T : V \to V$ in that basis. Consider also the matrices $B = (\langle Tv_i, v_j \rangle)$ and $G = (\langle v_i, v_j \rangle)$. Prove that $A^t = BG^{-1}$.

2 Let $f : M^2 \to \mathbf{R}^3$ be an isometric immersion of a surface, consider the frame of vector fields $\frac{\partial}{\partial x_1}$, $\frac{\partial}{\partial x_2}$ along f and the corresponding coefficients g_{ij} of the induced Riemannian metric.

a. Show that the coefficients of the second fundamental form of f are given by

$$b_{ij} = \det\left(\frac{\partial^2 f}{\partial x_i \partial x_j}, \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right) \cdot \det(g_{k\ell})^{-1/2},$$

with respect to some choice of unit normal vector field ξ.

[2] Globally flat normal bundle

[3] Normal bundle; normal connection; normal component of equation.

b. Deduce that the Gaussian curvature

$$K = \det A_\xi = \frac{\det(b_{ij})}{\det(g_{ij})}$$

and the mean curvature

$$H = \operatorname{tr} A_\xi = \frac{g_{11}b_{22} - 2g_{12}b_{12} + g_{22}b_{11}}{\det(g_{ij})}.$$

3 *a*. Let $\gamma : (a,b) \to \mathbf{R}^3$, $\xi : (a,b) \to S^2(1)$ be smooth curves. A parametrized surface of the form $f(u,v) = \gamma(u) + v\xi(v)$ is called a *ruled surface*. Investigate sufficient conditions for f to be an immersion. Compute that

$$K = \frac{-(\gamma' \cdot \xi')^2}{||(\gamma' + v\xi') \times \xi||^2}.$$

Deduce that the plane, cilinder and cone are flat surfaces.

b. For the *helicoid*

$$f(u,v) = (v\cos u, v\sin u, au)$$

($a > 0$), show that

$$K(u,v) = \frac{-a^2}{(a^2+v^2)^2} \tag{7.9.1}$$

and that it is a minimal surface. Deduce its principal curvatures. It is not difficult to show that the plane and the helicoid are the only complete ruled minimal surfaces in $\mathbf{R}^3$.

4 *a*. Let $\gamma : (a,b) \to \mathbf{R}^2$ be a smooth curve. A parametrized surface of the form $f(u,v) = (\gamma_1(v)\cos u, \gamma_1(v)\sin u, \gamma_2(v))$, where γ_1, γ_2 are the components of γ, is called a *surface of revolution*. Show that

$$K = \frac{\gamma_2'(\gamma_1'\gamma_2'' - \gamma_2'\gamma_1'')}{\gamma_1((\gamma_1')^2 + (\gamma_2')^2)^2)}.$$

In particular $K = -\gamma_1''/\gamma_1$ in case γ is parametrized by arc-length.

b. For the *torus of revolution*

$$f(u,v) = ((R + r\cos v)\cos u, (R + r\cos v)\sin u, r\sin v)$$

($R > r > 0$), show that

$$K = \frac{\cos v}{r(R + r\cos v)}.$$

c. For the *catenoid*

$$f(u,v) = (a\cosh(v/a)\cos u, a\cosh(v/a)\sin u, v)$$

($a > 0$), show that

$$K(u,v) = \frac{-1}{a^2\cosh^4(v/a)} \tag{7.9.2}$$

and that it is a minimal surface. It is not difficult to see that the only complete minimal surfaces of revolution in $\mathbf{R}^3$ are the plane and the catenoid. Interpret formulae (7.9.1) and (7.9.2) in view of exercise 2 of chapter 1.

5 Let M be a surface in $\mathbf{R}^3$ given as the pre-image of a regular value of a smooth map $f : U \to \mathbf{R}$, where U is an open subset of $\mathbf{R}^3$. Show that the second fundamental form of M is given by

$$B(u, v) = \frac{1}{||(\operatorname{grad} f)_p||} \operatorname{Hess}(f)(u, v)$$

for some choice of unit normal vector field, where $p \in M$ and u, $v \in T_pM$.

6 (The Beez-Killing theorem) *a.* Let S, $T : V \to V$ be self-adjoint linear operators on an Euclidean vector space V. Suppose that $\operatorname{rank}(S) \geq 3$ and $\Lambda^2 S = \Lambda^2 T : \Lambda^2 V \to \Lambda^2 V$. Prove that $S = \pm T$.

b. Let M be a (not necessarily complete) connected Riemannian manifold of dimension n and suppose $f : M \to \mathbf{R}^{n+1}$ is an isometric immersion such that the rank of the second fundamental form is at least 3 at every point. Prove that f is rigid.

7 Let $M \subset N \subset P$ be a chain of Riemannian submanifolds. Prove that if M is totally geodesic in N and N is totally geodesic in P, then M is totally geodesic in P.

8 Prove that each connected component of the fixed point set of an isometry of a Riemannian manifold is a properly embedded totally geodesic submanifold. Generalize the result to the fixed point set of a group of isometries.

9 Prove that the totally geodesic submanifolds of $\mathbf{R}P^n$ are the images of totally geodesic submanifolds of S^n under the projection $\pi : S^n \to \mathbf{R}P^n$. Deduce that the complete totally geodesic submanifolds of $\mathbf{R}P^n$ are isometric to $\mathbf{R}P^k$ for some $0 \leq k \leq n$; in particular, the cut-locus of a point in $\mathbf{R}P^n$ is a totally geodesic hypersurface isometric to $\mathbf{R}P^{n-1}$.

10 Consider the projection $\pi : S^{2n+1} \setminus \{0\} \to \mathbf{C}P^n$. Prove that there are exactly two kinds of complete totally geodesic submanifolds of $\mathbf{C}P^n$: (i) $\pi(V \cap S^{2n+1})$, where V is a complex subspace of $\mathbf{C}^{n+1}$; and (ii) $\pi(W \cap S^{2n+1})$, where W is a totally real subspace of $\mathbf{C}^{2n+1}$. Deduce that the complete totally geodesic submanifolds of $\mathbf{C}P^n$ are isometric to $\mathbf{C}P^k$ or to $\mathbf{R}P^k$ for some $0 \leq k \leq n$; in particular, the cut-locus of a point in $\mathbf{C}P^n$ is a totally geodesic submanifold isometric to $\mathbf{C}P^{n-1}$.

11 Let M^n be a Riemannian submanifold of $\mathbf{R}^{n+k}$. Fix a point $p \in M$ and a normal vector $\xi \in \nu_pM$. In this exercise we establish a canonical isomorphism $T_\xi(\nu M) \cong T_pM \oplus \nu_pM$.

a. Given $u \in T_pM$, consider a smooth curve $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$, $\gamma'(0) = u$ and take the parallel transport $\hat{\xi}$ of ξ along γ. Show that this defines a linear map $T_pM \to T_\xi(\nu M)$, and that this map is injective.

b. Given $\eta \in \nu_pM$, consider the line $s \mapsto \xi + s\eta$ in ν_pM. Show that it defines a linear map $\nu_pM \to T_\xi(\nu M)$, and that this map is injective.

c. Show that T_pM and ν_pM viewed as subspaces of $T_\xi(\nu M)$ meet only at 0. Deduce the above claim.

12 Let M^n be a Riemannian submanifold of $\bar{M} = \mathbf{R}^{n+k}$. Consider the normal exponential map $\exp^\perp : \nu M \to \mathbf{R}^{n+k}$ mapping $\xi \in \nu_pM$ to $p + \xi$.

a. Use exercise 11 to represent the differential $d(\exp^\perp)_\xi : T_pM \oplus \nu_pM \to T_pM \oplus \nu_pM$ as

$$\begin{pmatrix} \operatorname{id} - A_\xi & 0 \\ 0 & \operatorname{id} \end{pmatrix}.$$

b. Assume ξ is a unit vector and prove that $q = p + t\xi$ is a focal point of multiplicity m of M along the normal line through ξ if and only if $1/t$ is an eigenvalue of A_ξ of multiplicity m. Deduce that d is a focal distance of M along ξ if and only if $1/d$ is a principal curvature of A_ξ.

c. Generalize the above to other space forms to prove that: in S^{n+k}, d is a focal distance of M along ξ if and only if $\cot d$ is a principal curvature of A_ξ; in $\mathbf{R}H^{n+k}$, d is a focal distance of M along ξ if and only if $\coth d$ is a principal curvature of A_ξ.

d. In case $\bar{M} = S^{n+k}$, note that d is a focal distance of M along ξ if and only $\pi - d$ is a focal distance of M along $-\xi$.

13 (The Morse index theorem for submanifolds of Euclidean space) Let M be a Riemannian submanifold of $\bar{M} = \mathbf{R}^n$. For $q \in \mathbf{R}^n$, consider the square distance function

$$L_q : M \to \mathbf{R}, \qquad L_q(x) = \frac{1}{2}||x - q||^2.$$

a. Prove that $\mathrm{grad}(L_q)_p = (p - q)^\top$. Deduce that $p \in M$ is a critical point of L_q if and only if $v = q - p \in \nu_p M$.

b. Let $p \in M$ be a critical point of L_q and $v = q - p \in \nu_p M$. Prove that $\mathrm{Hess}(L_q)_p = I - A_v$ (exercise 13 of chapter 4).

c. The *nullity* of L_q at a critical point p is defined to be the nullity of the symmetric bilinear form $\mathrm{Hess}(L_q)_p$; such a critical point p is called *non-degenerate* if the nullity of L_q at p is zero. Use Exercise 12 to deduce that the nullity of L_q at a critical point p equals the multiplicity of q as a focal point of M along the geodesic segment $\overline{pq}$. Deduce that p is non-degenerate as a critical point of L_q if and only if q is a non-focal point of M along the geodesic segment $\overline{pq}$.

d. The *index* $\mathrm{ind}(L_q)_p$ of L_q at a critical point p is defined to be the index of the symmetric bilinear form $\mathrm{Hess}(L_q)_p$. Show that $\mathrm{ind}(L_q)_p = \sum_{t \in (0,1)} \ker(I - t\,A_v)$, where $v = q - p$. Combine this result with part (c) to deduce that $\mathrm{ind}(L_q)_p$ equals the sum of the multiplicities of $p + tv$ as a focal point to M for $t \in (0, 1)$.

e. Check that this result is a specialization of the Morse index theorem 7.5.4 to the case of Euclidean submanifolds.

14 Let M be a submanifold of a Riemannian manifold $\bar{M}$. Prove that the kth-osculating space $\mathcal{O}^k_p(M)$ of M at a point $p \in M$ is spanned by the k-th derivatives at 0 of all smooth curves $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$. (Hint: Consider the reparametrizations $\gamma(\vartheta(t))$ where ϑ is a polynomial function with $\vartheta(0) = 0$.)

15 Let M be a complete isoparametric submanifold of Euclidean space $\bar{M} = \mathbf{R}^n$. Fix a parallel normal vector field ξ along M. Consider $\pi_\xi : M \to M_\xi$ and let $\hat{p} \in M_\xi$. Prove that the connected components of the level set $\pi^{-1}(\hat{p})$ are compact isoparametric submanifolds of $\nu_{\hat{p}}(M_\xi)$, with curvature normals given exactly by those curvature normals v_i of M that satisfy $\langle \xi, v_i \rangle = 1$.

Bibliography

[Abr83] U. Abresch, *Isoparametric hypersurfaces with four or six principal curvatures*, Math. Ann. **264** (1983), 283–302.

[Ale04] M. M. Alexandrino, *Singular Riemannian foliations with sections*, Illinois J. Math. **48** (2004), no. 4, 1163–1182.

[Aub76] T. Aubin, *Équations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire*, J. Math. Pures Appl. (9) **55** (1976), no. 3, 269–296.

[Aub98] ______, *Some nonlinear problems in Riemannian geometry*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.

[BB76] L. Berard-Bergery, *Les variétés riemanniennes homogénes simplement connexes de dimension impaire à courbure strictement positive*, J. Math. Pures Appl. (9) **55** (1976), no. 1, 47–67.

[BBI01] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001.

[BCO16] J. Berndt, S. Console, and C. Olmos, *Submanifolds and holonomy*, second ed., Research Notes in Mathematics, vol. 434, Chapman & Hall/CRC, Boca Raton, 2016.

[Bel68] E. Beltrami, *Saggio di interpretazione della geometria non-euclidea*, Opere Mat. **1** (1868), 374–405.

[Bes78] A. Besse, *Manifolds all of whose geodesics are closed*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 93, Springer-Verlag, Berlin-New York, 1978, With appendices by D. B. A. Epstein, J.-P. Bourguignon, L. Bérard-Bergery, M. Berger and J. L. Kazdan.

[BI94] Y. Burago and S. Ivanov, *Riemannain tori without conjugate points are flat*, Geom. Anal. Funct. Anal. **5** (1994), 801–808.

[BK80] M. Berger and J. L. Kazdan, *A Sturm-Liouville inequality with applications to an isoperimetric inequality for volume, injectivity radius and wiedersehens manifolds*, General inequalities (E. F. Beckenbach, ed.), vol. II, Birkhäuser, Basel, 1980, pp. 367–377.

[Bla30] W. Blaschke, *Vorlesungen über Differentialgeometrie*, 3rd ed., Springer, Berlin, 1930.

[BR96] A. Bellaïche and J.-J. Risler (eds.), *Sub-Riemannian geometry*, Progress in Mathematics, vol. 144, Birkhäuser Verlag, Basel, 1996.

[Car38] É. Cartan, *Familles de surfaces isoparamétriques dans les espaces à courbure constante*, Ann. Mat. Pura Appl. **17** (1938), no. 1, 177–191, Also in: *Ouevres Complètes*, Partie III, Vol. 2, 1431–1445.

[Car39a] ______, *Sur des familles remarquables d'hypersurfaces isoparamétriques dans les espaces sphériques*, Math. Z. **45** (1939), 335–367, Also in: *Ouevres Complètes*, Partie III, Vol. 2, 1447–1479.

[Car39b] ______, *Sur quelques familles remarquables d'hypersurfaces*, C. R. Congrès Math. Liege (1939), 30–41, Also in: *Ouevres Complètes*, Partie III, Vol. 2, 1481–1492.

[Car40] ______, *Sur des familles d'hypersurfaces isoparamétriques des espaces sphériques à 5 et à 9 dimensions*, Univ. Nac. Tucumán, Revista A (1940), no. 1, 5–22, Also in: *Ouevres Complètes*, Partie III, Vol. 2, 1513–1530.

[Car76] M. P. do Carmo, *Differential geometry of curves and surfaces*, Prentice-Hall, Englewood Cliffs, NJ, 1976.

[CCJ07] T. E. Cecil, Q. S. Chi, and G. R. Jensen, *Isoparametric hypersurfaces with four principal curvatures*, Ann. Math. (2) **166** (2007), no. 1, 1–76.

[Cha86] C. S. Charlap, *Bieberbach Groups and Flat Manifolds*, Springer-Verlag, New York, 1986.

[Chi12] Q. S. Chi, *A new look at condition A*, Osaka J. Math. **49** (2012), no. 1, 133–166.

[CM11] T. H. Colding and W. P. Minicozzi, II, *A course in minimal surfaces*, Graduate Studies in Mathematics, vol. 121, American Mathematical Society, Providence, RI, 2011.

[Dad85] J. Dadok, *Polar actions induced by actions of compact Lie groups*, Trans. Amer. Math. Soc. **288** (1985), 125–137.

[DN85] J. Dorfmeister and E. Neher, *Isoparametric hypersurfaces, case $g = 6$, $m = 1$*, Comm. Algebra **13** (1985), no. 11, 2299–2368.

[dR73] G. de Rham, *Variétés Différentiables*, third ed., Hermann, Paris, 1973.

[EBG69] E. de Giorgi E. Bombieri and E. Giusti, *Minimal cones and the Bernstein problem*, Invent. Math. **7** (1969), 243–268.

[EE69] C. J. Earle and J. Eells, *A fibre bundle description of Teichmüller theory*, J. Differential Geom. **3** (1969), 19–43.

[Ehr51] C. Ehresmann, *Les connexions infinitésimales dans un espace fibré différentiable*, Colloque de topologie (espaces fibrés), Bruxelles, 1950, Georges Thone, Liége; Masson et Cie., Paris, 1951, pp. 29–55.

[Eps75] D. B. A. Epstein, *Natural tensors on Riemannian manifolds*, J. Differential Geometry **10** (1975), no. 4, 631–645.

[FKM81] D. Ferus, H. Karcher, and H.E. Münzner, *Cliffordalgebren und neue isoparametrische Hyperflächen*, Math. Z. **177** (1981), 479–502.

[GH12] C. Gorodski and E. Heintze, *Homogeneous structures and rigidity of isoparametric submanifolds in hilbert space*, J. Fixed Point Theory Appl. **11** (2012), no. 1, 93–136.

[Gre63] L. W. Green, *Aufwiedersehensflächen*, Ann. Math. (2) **78** (1963), 289–299.

[Ham82] R. Hamilton, *Three-manifolds with positive Ricci curvature*, J. Differential Geom. **17** (1982), 255–306.

[Hel78] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Mathematics, vol. 80, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.

[Hil01] D. Hilbert, *Über Flächen von constanter Gaussscher Krümmung*, Trans. Amer. Math. Soc. **2** (1901), no. 1, 87–99.

[HL71] W.-Y. Hsiang and H. B. Lawson Jr., *Minimal submanifolds of low cohomogeneity*, J. Differential Geom. **5** (1971), no. 1-2, 1–38.

[HL99] E. Heintze and X. Liu, *Homogeneity of infinite dimensional isoparametric submanifolds*, Ann. of Math. **149** (1999), 149–181.

[Hop48] E. Hopf, *Closed surfaces without conjugate points*, Proc. Nat. Acad. Sci., USA **34** (1948), 47–51.

[Hop89] H. Hopf, *Differential geometry in the large*, second ed., Lecture Notes in Mathematics, vol. 1000, Springer-Verlag, Berlin, 1989, notes taken by Peter Lax and John W. Gray, with a preface by S. S. Chern.

[HR31] H. Hopf and W. Rinow, *Über den Begriff der vollständigen differentialgeometrischen fläche*, Comment. Math. Helv. **3** (1931), no. 1, 209–225.

[Imm08] S. Immervoll, *On the classification of isoparametric hypersurfaces with four distinct principal curvatures in spheres*, Ann. Math. (2) **168** (2008), no. 3, 1011–1024.

[Jos06] J. Jost, *Compact Riemann surfaces. an introduction to contemporary mathematics*, 3rd ed., Universitext, Springer-Verlag, Berlin, 2006.

[Kli95] W. Klingenberg, *Riemannian geometry*, 2nd ed., de Gruyter, Berlin, 1995.

[KN96] S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, Wiley Classics Library, vol. I, John Wiley & Sons, Inc., New York, 1996, Reprint of the 1963 original.

[Kul70] R. S. Kulkarni, *Curvature and metric*, Ann. of Math. (1970), 311–331.

[Law80] H. B. Lawson Jr., *Lectures on minimal submanifolds*, second ed., vol. I, Mathematics Lecture Series, no. 9, Publish or Perish, Inc., Wilmington, Del., 1980.

[LC37] T. Levi-Cività, *Famiglie di superficie isoparametrische nellórdinario spazio euclideo*, Rend. Acc. Naz. Lincei **XXVI** (1937), 355–362, Also in: *Opere matematici, **V**, 657–664.*

[Lev05] T. Levi-Cività, *The absolute differential calculus: Calculus of tensors*, Dover Phoenix Editions, Dover publications, 2005, Translated from the Italian original (*Lezioni di calcolo differenziale assoluto*, 1925) by Marjorie Long.

[LF51] L. A. Lyusternik and A. I. Fet, *Variational problems on closed manifolds*, Doklady Akad. Nauk SSSR (N.S.) **81** (1951), 17–18.

[LŠ47] L. Lyusternik and L. Šnirel'man, *Topological methods in variational problems and their application to the differential geometry of surfaces*, Uspehi Matem. Nauk (N.S.) **2** (1947), no. 1(17), 166–217.

[Mil97] J. W. Milnor, *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.

[Miy09] R. Miyaoka, *The Dorfmeister-Neher theorem on isoparametric hypersurfaces*, Osaka. J. Math. **46** (2009), 695–715.

[Miy13] ______, *Isoparametric hypersurfaces with* $(g, m) = (6, 2)$, Ann. of Math. (2) **177** (2013), 53–110.

[Miy15] ______, *Remarks on the Dorfmeister-Neher theorem on isoparametric hypersurfaces*, Osaka J. Math. **52** (2015), 373–377.

[Mon02] R. Montgomery, *A tour of sub-riemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, vol. 91, American Mathematical Society, Providence, RI, 2002.

[MS39] S. B. Myers and N. E. Steenrod, *The group of isometries of a Riemannian manifold*, Ann. of Math. (2) **40** (1939), no. 2, 400–416.

[MT06] J. W. Morgan and G. Tian, *Ricci Flow and the Poincaré Conjecture*, arXiv:math/0607607v2 [math.DG], 2006.

[Mue80] H. F. Muenzner, *Isoparametrische Hyperflaechen in Sphaeren, I*, Math. Ann. **251** (1980), 57–71.

[Mue81] ______, *Isoparametrische Hyperflaechen in Sphaeren, II*, Math. Ann. **256** (1981), 215–232.

[Mun84] J. R. Munkres, *Elements of algebraic topology*, Addison-Wesley Publishing Company, Menlo Park, CA, 1984.

[Nas56] J. F. Nash, *The imbedding problem for Riemannian manifold*, Ann. of Math. (2) **63** (1956), 20–63.

[Nom54] K. Nomizu, *Invariant affine connections on homogeneous spaces*, Amer. J. Math. **76** (1954), 33–65.

[Nom73] ———, *Some results in E. Cartan's theory of isoparametric families of hypersurfaces*, Bull. Amer. Math. Soc. **79** (1973), 1184–1188.

[Olm93] C. Olmos, *Isoparametric submanifolds and their homogeneous structures*, J. Differential Geom. **38** (1993), 225–234.

[Oss86] R. Osserman, *A survey of minimal surfaces*, second ed., Dover Publications, Inc., New York, 1986.

[OT75] H. Ozeki and M. Takeuchi, *On some types ofisoparametric hypersurfaces in spheres. I*, Tôhoku Math. J. **27** (1975), 515–559.

[PT87] R. S. Palais and C.-L. Terng, *A general theory of canonical forms*, Trans. Amer. Math. Soc. **300** (1987), 771–789.

[PT88] ———, *Critical point theory and submanifold geometry*, Lecture Notes in Mathematics, vol. 1353, Springer-Verlag, Berlin, 1988.

[Rad24] T. Radó, *Über den Begriff der Riemannsche Fläche*, Acta Univ. Szeged **2** (1924), 101–121.

[Rez94] A. G. Reznikov, *The weak Blaschke conjecture for* $C\mathrm{P}^n$, Invent. Math. **117** (1994), no. 3, 447–454.

[Rie53] B. Riemann, *Über die Hypothesen, welche der Geometrie zur Grunde liegen (Habilitationschrift 1854)*, pp. 272–287, Dover Publ. Inc., New York, 1953, Reprint of 1892 edition and 1902 supplement.

[RL00] M. M. G. Ricci and T. Levi-Cività, *Méthodes de calcul différentiel absolu et leurs applications*, Math. Ann. **54** (1900), no. 1-2, 125–201.

[Sch84] R. Schoen, *Conformal deformation of a Riemannian metric to constant scalar curvature*, J. Differential Geom. **20** (1984), 479–495.

[Seg38] B. Segre, *Famiglie di ipersuperficie isoparametrische negli spazi euclidei ad un qualunque numero di dimensioni*, Rend. Acc. Naz. Lincei **XXVII** (1938), 203–207.

[Sif16] A. Siffert, *Classification of isoparametric hypersurfaces in spheres with* $(g, m) = (6, 1)$, Proc. Amer. Math. Soc. **144** (2016), no. 5, 2217–2230.

[Sim83] L. Simon, *Survey lectures on minimal submanifolds*, Ann. of Math. Stud., vol. 103, pp. 3–52, Princeton Univ. Press, Princeton, NJ, 1983.

[Spi70] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Boston, 1970.

[Sto99] S. Stolz, *Multiplicities of Dupin hypersurfaces*, Invent. Math. **138** (1999), no. 2, 253–279.

[Str34] D. J. Struik, *Theory of linear connections*, Ergebnisse d. Math., vol. 3, Springer, 1934.

[Ter85] C.-L. Terng, *Isoparametric submanifolds and their Coxeter groups*, J. Differential Geom. **21** (1985), 79–107.

[Ter89] ———, *Proper Fredholm submanifolds of Hilbert space*, J. Differential Geom. **29** (1989), 9–47.

[Thu97] W. P. Thurston, *Three-dimensional Geometry and Topology*, vol. 1, Princeton Mathematical Series, no. 35, Princeton University Press, Princeton, NJ, 1997.

[TT72] R. Takagi and T. Takahashi, *On the principle curvatures of homogeneous hypersurfaces in a sphere*, pp. 469–481, Kinokuniya, Tokyo, 1972.

[TT95] C.-L. Terng and G. Thorbergsson, *Submanifold geometry in symmetric spaces*, J. Differential Geom. **42** (1995), no. 3, 665–718.

[Wal72] N. R. Wallach, *Compact homogeneous Riemannian manifolds with strictly positive curvature*, Ann. of Math. (2) **96** (1972), 277–295.

[Wal97] G. Walschap, *Cut and conjugate loci in two-step nilpotent Lie groups*, J. Geom. Anal. **7** (1997), no. 2, 343–355.

[War67] F. W. Warner, *Conjugate loci of constant order*, Ann. Math. (2) **86** (1967), no. 1, 192–212.

[War83] ———, *Foundations of Differentiable Manifolds and Lie Groups*, Graduate Texts in Mathematics, vol. 94, Springer-Verlag, New York, 1983.

[Wei74] A. Weinstein, *On the volume of manifolds all of whose geodesics are closed*, J. Differential Geom. **9** (1974), 513–517.

[Wel08] R. O. Wells, Jr., *Differential analysis on complex manifolds*, third edition ed., Graduate Texts in Mathematics, vol. 65, Springer, New York, 2008, With a new appendix by Oscar Garcia-Prada.

[Wol84] J. A. Wolf, *Spaces of Constant Curvature*, fifth ed., Publish or Perish, 1984.

[Yam60] H. Yamabe, *On a deformation of riemannian structures on compact manifolds*, Osaka Math. J. **12** (1960), 21–37.

[Yan80] C. T. Yang, *Odd dimensional wiedersehens manifolds are spheres*, J. Differential Geom. **15** (1980), 91–96.

[Zil07] W. Ziller, *Examples of Riemannian manifolds with non-negative sectional curvature*, arXiv:math/0701389 [math.DG], 2007.

Index