Lecture Notes on

# Riemannian Geometry

Urs Lang

ETH Zürich

Spring Semester 2020

**Preliminary and incomplete version**

June 16, 2020

**Acknowledgements**

These notes are based on the lecture courses *Differential Geometry II* I taught at ETH Zurich in the Spring Semesters 2017 and 2020. I thank Alessandro Fasse for his help in typing the first version of these notes.

# Contents


**1 Riemannian manifolds** **1**
Riemannian metrics and distance functions . . . 1
The Levi-Civita connection . . . 3
Vector fields along maps . . . 7
Geodesics and the exponential map . . . 10
The Hopf–Rinow Theorem . . . 14
Geodesic metric spaces . . . 16

**2 Curvature** **19**
The curvature tensor . . . 19
Sectional curvature . . . 22
Ricci and scalar curvature . . . 23
Curvature of submanifolds . . . 29
Riemannian products . . . 31

**3 Jacobi Fields** **33**
Second variation of arc length . . . 33
Jacobi fields . . . 35
Conjugate points . . . 37
The Rauch Comparison Theorem . . . 40
Focal points and the Rauch–Berger Theorem . . . 42

**4 Riemannian Submersions and Coverings** **45**
Riemannian submersions . . . 45
Curvature of Riemannian submersions . . . 47
Riemannian coverings and space forms . . . 50
Hadamard manifolds . . . 54
Isometries of Hadamard manifolds . . . 55

**5 Triangle comparison** **59**
Some model space geometry . . . 59
Alexandrov comparisons . . . 61

Toponogov's Theorem . . . . . . . . . . . . . . . . . . . . . . . . . . 64
Open manifolds of non-negative curvature . . . . . . . . . . . . . . . . 67

**6 Volume comparison and applications** **75**
Volume comparison theorems . . . . . . . . . . . . . . . . . . . . . . . . 75
Growth of the fundamental group . . . . . . . . . . . . . . . . . . . . . 80
Gromov–Hausdorff convergence . . . . . . . . . . . . . . . . . . . . . . 83
Diffeomorphism finiteness . . . . . . . . . . . . . . . . . . . . . . . . . 87

**Bibliography** **89**

# Chapter 1

# Riemannian manifolds

## Riemannian metrics and distance functions

In the following, $M$ will always denote an $m$-dimensional smooth manifold, that is, a differentiable manifold of class $C^\infty$. The tangent and cotangent bundles of $M$ are denoted by $TM$ and $TM^*$, respectively, the vector spaces of $C^\infty$ sections by $\Gamma(TM)$ (vector fields) and $\Gamma(TM^*)$ (1-forms). For integers $r, s \geq 0$,

$$T_{r,s}M = \underbrace{TM \otimes \ldots \otimes TM}_{r} \otimes \underbrace{TM^* \otimes \ldots \otimes TM^*}_{s}$$

denotes the $(r, s)$-tensor bundle of $M$; thus $T_{1,0}M = TM$, $T_{0,1}M = TM^*$, and, by convention, $T_{0,0}M = C^\infty(M)$. For $p \in M$, the fiber

$$T_{r,s}M_p = TM_p \otimes \ldots \otimes TM_p \otimes TM_p^* \otimes \ldots \otimes TM_p^*$$

will be identified with the vector space of multilinear functions

$$\underbrace{TM_p^* \times \ldots \times TM_p^*}_{r} \times \underbrace{TM_p \times \ldots \times TM_p}_{s} \to \mathbb{R}.$$

With this identification, an $(r, s)$-tensor field $T \in \Gamma(T_{r,s}M)$ corresponds to an $\mathbb{R}$-multilinear map

$$T \colon (\Gamma(TM^*))^r \times (\Gamma(TM))^s \to C^\infty(M)$$

that is $C^\infty(M)$-homogeneous in every argument. A $(1, s)$-tensor field $T$ is often viewed as the $s$-linear map $(\Gamma(TM))^s \to \Gamma(TM)$, denoted again by $T$, such that $T(\theta, V_1, \ldots, V_s) = \theta(T(V_1, \ldots, V_s))$ for all $\theta \in \Gamma(TM^*)$ and $V_1, \ldots, V_s \in \Gamma(TM)$.

**1.1 Definition** A *Riemannian metric* $g$ on $M$ is a $(0, 2)$-tensor field (that is, $g \in \Gamma(T_{0,2}M)$) such that for all $p \in M$,

$$g_p \colon TM_p \times TM_p \to \mathbb{R}$$

is an inner product (a positive definite symmetric bilinear form). The pair $(M, g)$ is called a *Riemannian manifold*.

For any chart $\varphi\colon U \to \varphi(U) \subset \mathbb{R}^m$ of $M$, the restriction of $g$ to $U$ has a representation of the form

$$g|_U = \sum_{i,j=1}^{m} g_{ij}\, d\varphi^i \otimes d\varphi^j,$$

where $(d\varphi^i \otimes d\varphi^j)(X, Y) = d\varphi^i(X)d\varphi^j(Y) = X(\varphi^i)Y(\varphi^j)$ for $X, Y \in \Gamma(TU)$. Note that $g_{ij} = g\big(\frac{\partial}{\partial\varphi^i}, \frac{\partial}{\partial\varphi^j}\big) = g_{ji} \in C^\infty(U)$.

It is customary to just write $\langle\,\cdot\,,\cdot\,\rangle$ in place of $g$.

**1.2 Remark** On every smooth manifold $M$ there exists a Riemannian metric $g$. This can be shown by means of a partition of unity (exercise).

If $(\bar{M}, \bar{g})$ is a Riemannian manifold and $F\colon M \to \bar{M}$ is an immersion of another smooth manifold $M$, then

$$(F^*\bar{g})_p(v, w) := \bar{g}_{F(p)}(F_*v, F_*w) = \bar{g}_{F(p)}(dF_p(v), dF_p(w))$$

$v, w \in TM_p$, defines the *pull-back metric* $F^*\bar{g}$ on $M$.

**1.3 Definition** Two Riemannian manifolds $(M, g)$ and $(\bar{M}, \bar{g})$ are called *isometric* if there exists a diffeomorphism $F\colon M \to \bar{M}$ such that $F^*\bar{g} = g$; then $F$ is called an *isometry* from $(M, g)$ to $(\bar{M}, \bar{g})$. A smooth map $F\colon M \to \bar{M}$ is called an *isometric immersion* if $F^*\bar{g} = g$. If, in addition, $F$ is an embedding, then $F$ is called an *isometric embedding*.

By the famous Nash embedding theorem [Na1956], every $C^\infty$ Riemannian manifold $(M, g)$ of dimension $m$ admits an isometric $C^\infty$ embedding into the Euclidean space $(\mathbb{R}^n, \langle\,\cdot\,,\cdot\,\rangle_{\mathrm{eucl}})$ for some $n = n(m)$ (a quadratic polynomial in $m$). The (much more flexible) $C^1$ case was considered earlier by Nash [Na1954] and Kuiper [Ku1955].

We now prove that a connected Riemannian manifold is, in a natural way, also a metric space.

**1.4 Theorem (distance function)** *Let $(M, g)$ be a connected Riemannian manifold. For every pair of points $p, q \in M$, define $d(p, q)$ as the infimum of $L(c)$ taken over all piecewise $C^1$ curves $c\colon [a, b] \to M$ from $p$ to $q$, that is, with $c(a) = p$ and $c(b) = q$. This yields a distance function (metric) on $M$, and the topology induced by $d$ agrees with the given topology of $M$.*

For non-connected Riemannian manifolds, the result still holds, except that $d(p, q) = \infty$ for points in distinct connected components.

*Proof*: Since $M$ is connected, any two points in $M$ can be connected by a piecewise $C^1$ curve $c\colon [a,b] \to M$, and any such curve has finite length. Hence $d$ is finite, and clearly $d \geq 0$, $d(p,p) = 0$, $d(p,q) = d(q,p)$, and $d(p,q) \leq d(p,p') + d(p',q)$ for all $p, p', q \in M$. To show that $d$ is a metric, it remains to check that $d(p,q) > 0$ if $p \neq q$. Given such $p$ and $q$, let $(\varphi, U)$ be a chart of $M$ with $p \in U$ and $\varphi(p) = 0$. Let $\epsilon > 0$ be such that the ball $\bar{B}_\epsilon := \{x \in \mathbb{R}^m : |x|_{\mathrm{eucl}} \leq \epsilon\}$ is contained in $\varphi(U)$, and consider the compact set $K := \varphi^{-1}(\bar{B}_\epsilon)$. Since $M$ is Hausdorff, $K$ is closed in $M$, and we can fix $\epsilon > 0$ so that $q \notin \varphi^{-1}(B_\epsilon)$, where $B_\epsilon$ is the open ball. Let now $c\colon [a,b] \to M$ be a piecewise $C^1$ curve from $p$ to $q$, and put

$$s := \sup\{t \in [a,b] : c([a,t]) \subset K\}.$$

Then $c([a,s]) \subset K$ since $K$ is closed, and $c(s) \notin \varphi^{-1}(B_\epsilon)$. Hence, $\bar{c} := \varphi \circ c|_{[a,s]}$ is a piecewise $C^1$ curve in $\bar{B}_\epsilon$ connecting 0 to a boundary point of $\bar{B}_\epsilon$. Put $\bar{g} := (\varphi^{-1})^* g$ on $\varphi(U)$. Since $\bar{B}_\epsilon$ is compact, there exists a constant $\lambda > 0$ such that $|v|_{\bar{g}_x} \geq \lambda |v|_{\mathrm{eucl}}$ for all $v \in \mathbb{R}^m$ and $x \in \bar{B}_\epsilon$. Now $L(c) \geq L(c|_{[a,s]}) = L_{\bar{g}}(\bar{c}) \geq \lambda\, L_{\mathrm{eucl}}(\bar{c}) \geq \lambda\epsilon$, independently of the choice of $c$, thus $d(p,q) \geq \lambda\epsilon > 0$.

The same argument shows that if $p \in M$, $(\varphi, U)$ is a chart with $\varphi(p) = 0$, and $\epsilon > 0$ is such that $\bar{B}_\epsilon \subset \varphi(U)$, then there is a $\delta > 0$ such that

$$B(p,\delta) := \{p' \in M : d(p,p') < \delta\} \subset \varphi^{-1}(B_\epsilon).$$

Furthermore, there is a constant $\delta' \in (0,\epsilon]$ such that the line segment from 0 to any point in $B_{\delta'}$ has length $< \epsilon$ with respect to $\bar{g} = (\varphi^{-1})^* g$, thus

$$\varphi^{-1}(B_{\delta'}) \subset B(p,\epsilon).$$

These inclusions show that the two topologies agree. □

## The Levi-Civita connection

**1.5 Definition** Let $M$ be a smooth $m$-dimensional manifold. A *connection* $\nabla$ on $TM$ is an $\mathbb{R}$-bilinear map $\nabla\colon \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$, where $\nabla(X,Y)$ is written as $\nabla_X Y$, with the following properties:

(1) $\nabla_{fX} Y = f \nabla_X Y$,

(2) $\nabla_X (fY) = X(f)Y + f\nabla_X Y$,

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

Thus, a connection $\nabla$ is $C^\infty(M)$-homogeneous (“tensorial”) in the first argument and satisfies a product derivation rule in the second. For a given vector field $Y \in \Gamma(TM)$, the map $\nabla_. Y\colon \Gamma(TM) \to \Gamma(TM)$ defines a $(1,1)$-tensor field, thus

$(\nabla_X Y)_p$ depends only on $X_p$. On the other hand, if $X$ is fixed, $(\nabla_X Y)_p$ depends only on the restriction of $Y$ to a neighborhood $U$ of $p$. For this, it suffices to check that $(\nabla_X Y)_p = 0$ if $Y|_U \equiv 0$. Choose $f \in C^\infty(M)$ such that $f(p) = 0$ and $f|_{M\setminus U} \equiv 1$. Then $Y = fY$, hence $\nabla_X Y = X(f)Y + f\nabla_X Y$ by (2). Since $Y_p = 0$ and $f(p) = 0$, this shows that $(\nabla_X Y)_p = 0$. In particular, for every open set $U \subset M$ there is a well-defined induced connection on $TU$, still denoted by $\nabla$, such that

$$\nabla_{X|_U}(Y|_U) = (\nabla_X Y)|_U$$

for all $X, Y \in \Gamma(TM)$. In fact, more is true, see Remark 1.7 below.

We now consider the representation of a connection $\nabla$ in local coordinates or, more generally, with respect to a *moving frame* $(A_1, \dots, A_m)$ on an open set $U \subset M$. This means that $A_i \in \Gamma(TU)$ for all $i$ and $(A_1|_p, \dots, A_m|_p)$ is a basis of $TM_p$ for all $p \in U$. (For example, $A_i = \frac{\partial}{\partial\varphi^i}$ for a chart $(\varphi, U)$.) The *Christoffel symbols* of $\nabla$ with respect to $(A_1, \dots, A_m)$ are defined by

$$\nabla_{A_i} A_j = \sum_{k=1}^{m} \Gamma_{ij}^k A_k,$$

where $\Gamma_{ij}^k \in C^\infty(U)$ for all $i, j, k \in \{1, \dots, m\}$.

**1.6 Lemma** *For vector fields $X = \sum_{i=1}^m X^i A_i$ and $Y = \sum_{j=1}^m Y^j A_j$ on $U$,*

$$\nabla_X Y = \sum_{k=1}^{m} \biggl( X(Y^k) + \sum_{i,j=1}^{m} X^i Y^j \Gamma_{ij}^k \biggr) A_k.$$

*Proof*: By property (2) in Definition 1.5,

$$\nabla_X Y = \sum_{j=1}^{m} \nabla_X (Y^j A_j) = \sum_{j=1}^{m} \bigl( X(Y^j) A_j + Y^j \, \nabla_X A_j \bigr).$$

Furthermore, by property (1) and the definition of the Christoffel symbols,

$$\nabla_X A_j = \sum_{i=1}^{m} \nabla_{X^i A_i} A_j = \sum_{i=1}^{m} X^i \, \nabla_{A_i} A_j = \sum_{i,k=1}^{m} X^i \Gamma_{ij}^k A_k.$$

Combining these identities we obtain the result. □

This can also be written as follows. Let $(\alpha^1, \dots, \alpha^m)$ be the *coframe field* dual to $(A_1, \dots, A_m)$, that is, $\alpha^j \in \Gamma(TU^*)$ and $\alpha^j(A_i) = \delta_j^i$ for all $i, j \in \{1, \dots, m\}$. (For example, $A_i = \frac{\partial}{\partial\varphi^i}$ and $\alpha^j = \partial\varphi^j$.) The 1-forms

$$\omega_j^k := \sum_{i=1}^{m} \Gamma_{ij}^k \alpha^i$$

for $j,k \in \{1,\dots,m\}$ are called *connection forms* of $\nabla$ with respect to the given frame field $(A_1,\dots,A_m)$. For $X$ and $Y$ as above, $\alpha^i(X) = X^i$, thus

$$\nabla_X Y = \sum_{k=1}^{m} \Bigl( dY^k(X) + \sum_{j=1}^{m} Y^j \omega_j^k(X) \Bigr) A_k$$

or, more briefly, $\nabla Y = \sum_k \bigl( dY^k + \sum_j Y^j \omega_j^k \bigr) A_k$.

**1.7 Remark** From Lemma 1.6 we see that $(\nabla_X Y)_p$ depends in fact only on $X_p$ and the values of $Y$ along a curve $c\colon (-\epsilon,\epsilon) \to M$ with $c(0) = p$ and $c'(0) = X_p$, because $X_p(Y^k) = dY_p^k(X_p) = (Y^k \circ c)'(0)$.

**1.8 Definition** Let $M$ be a smooth manifold with a connection $\nabla$ on $TM$.

(1) The map $T\colon \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$,

$$T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y],$$

is called the *torsion tensor* of $\nabla$, and $\nabla$ is *torsion-free* if $T \equiv 0$.

(2) Given a Riemannian metric $g = \langle\,\cdot\,,\cdot\,\rangle$ on $M$, the connection $\nabla$ is said to be *compatible* with $g$ if, for all $X,Y,Z \in \Gamma(TM)$,

$$Z\langle X,Y\rangle = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle,$$

where $Z_p$ acts as derivation on the function $q \mapsto \langle X_q, Y_q\rangle$.

Note that, for any $f \in C^\infty(M)$,

$$\begin{aligned} T(fX,Y) &= f\nabla_X Y - \bigl(Y(f)X + f\nabla_Y X\bigr) - \bigl(f[X,Y] - Y(f)X\bigr) \\ &= fT(X,Y) \end{aligned}$$

and, hence, $T(X,fY) = -T(fY,X) = -fT(Y,X) = fT(X,Y)$. Thus the map $T$ defines a $(1,2)$-tensor field on $M$.

**1.9 Theorem (Levi-Civita connection)** *For every Riemannian manifold $(M,g)$ there exists a unique connection on $TM$ that is torsion-free and compatible with $g$.*

This connection, which we denote by $D$ instead of $\nabla$, is called the *Levi-Civita connection* of $g = \langle\,\cdot\,,\cdot\,\rangle$ and is characterized by the property that the *Koszul formula*

$$\text{(KF)} \qquad \begin{aligned} 2\langle D_X Y, Z\rangle = {}& X\langle Y,Z\rangle + Y\langle X,Z\rangle - Z\langle X,Y\rangle \\ & - \langle X,[Y,Z]\rangle - \langle Y,[X,Z]\rangle + \langle Z,[X,Y]\rangle \end{aligned}$$

holds for all $X,Y,Z \in \Gamma(TM)$.

*Proof*: Suppose first that there exists a connection $D$ on $TM$ that is compatible with $g = \langle\,\cdot\,,\cdot\,\rangle$ and torsion-free. Then, starting with the first three terms on the right-hand side of (KF) and using these two properties, one finds that

$$\begin{aligned}
&X\langle Y,Z\rangle + Y\langle X,Z\rangle - Z\langle X,Y\rangle \\
&\quad = \langle X, D_Y Z - D_Z Y\rangle + \langle Y, D_X Z - D_Z X\rangle + \langle Z, D_X Y + D_Y X\rangle \\
&\quad = \langle X, [Y,Z]\rangle + \langle Y, [X,Z]\rangle + \langle Z, 2D_X Y - [X,Y]\rangle.
\end{aligned}$$

Thus (KF) holds for $D$, for all $X,Y,Z \in \Gamma(TM)$, and this determines $D$ uniquely. For the existence, it remains to verify that the unique $\mathbb{R}$-bilinear map $D\colon \Gamma(TM)\times\Gamma(TM) \to \Gamma(TM)$ given by (KF) has the desired properties.

To see that $D_{fX}Y = fD_XY$, replace $X$ by $fX$ in (KF). Note that $f$ can be factored out of the sum of the second and the last term on the right-hand side, because $Y\langle fX,Z\rangle = Y(f)\langle X,Z\rangle + fY\langle X,Z\rangle$ and $\langle Z,[fX,Y]\rangle = \langle Z, f[X,Y] - Y(f)X\rangle$. The same holds for the third and fifth term, and the result follows.

Next, replace $Y$ by $fY$ in (KF). Then, as above, $f$ can be factored out of the term $-Z\langle X,fY\rangle - \langle X,[fY,Z]\rangle$, whereas for the first and last term on the right,

$$X\langle fY,Z\rangle + \langle Z,[X,fY]\rangle = f\big(X\langle Y,Z\rangle + \langle Z,[X,Y]\rangle\big) + 2\langle X(f)Y,Z\rangle.$$

It follows that $2\langle D_X(fY),Z\rangle = 2\langle X(f)Y,Z\rangle + f\cdot 2\langle D_XY,Z\rangle$, which yields the product rule $D_X(fY) = X(f)Y + fD_XY$.

To verify that $D$ is torsion-free, note that the expression consisting of the first five terms on the right-hand side of (KF) is symmetric in $X$ and $Y$. Thus

$$2\langle D_XY - D_YX, Z\rangle = \langle Z, [X,Y]-[Y,X]\rangle = 2\langle Z,[X,Y]\rangle.$$

Similarly, to check that $D$ is compatible with $g$, note that switching $Y$ and $Z$ in (KF) does not affect the first term on the right-hand side but corresponds to a sign change for the remaining part. Thus

$$2\langle D_XY,Z\rangle + 2\langle D_XZ,Y\rangle = 2X\langle Y,Z\rangle.$$

This concludes the proof. $\square$

If $(\varphi,U)$ is a chart of $M$ and $A_i := \frac{\partial}{\partial\varphi^i}$ are the coordinate vector fields, then $[A_i,A_j] = 0$, hence the Koszul formula yields

$$\sum_{k=1}^m \Gamma_{ij}^k g_{kl} = \langle D_{A_i}A_j, A_l\rangle = \frac{1}{2}\big(A_i g_{jl} + A_j g_{il} - A_l g_{ij}\big).$$

Solving this equation for $\Gamma_{ij}^k$, we obtain the expression

$$\Gamma_{ij}^k = \frac{1}{2}\sum_{l=1}^m g^{kl}\left(\frac{\partial g_{jl}}{\partial\varphi^i} + \frac{\partial g_{il}}{\partial\varphi^j} - \frac{\partial g_{ij}}{\partial\varphi^l}\right),$$

where $(g^{kl})$ denotes the matrix inverse to $(g_{kl})$. Note that $\Gamma_{ij}^k = \Gamma_{ji}^k$.

## Vector fields along maps

Let $F\colon N \to M$ be a smooth map. By a *vector field $Y$ along $F$* we mean a smooth map $Y\colon N \to TM$ such that $Y_p := Y(p) \in TM_{F(p)}$ for all $p \in N$. We write $\Gamma(F^*TM)$ for the set of all vector fields along $F$. For example, if $Z \in \Gamma(TM)$, then $Z \circ F \in \Gamma(F^*TM)$. On the other hand, any $X \in \Gamma(TN)$ induces a vector field

$$F_*X \in \Gamma(F^*TM), \quad F_*X_p := dF_p(X_p).$$

An important special case of this is the velocity vector field

$$c' := \frac{d}{dt}c := c_*\Big(\frac{d}{dt}\Big) \in \Gamma(c^*TM)$$

of a curve $c\colon I \to M$ with parameter $t$.

**1.10 Proposition** *Let $F\colon N \to M$ be a smooth map, and let $\nabla$ be a connection on $TM$. Then there is a unique $\mathbb{R}$-bilinear map $\nabla^F : \Gamma(TN)\times\Gamma(F^*TM) \to \Gamma(F^*TM)$, where $\nabla^F(X,Y)$ is written as $\nabla^F_X Y$, with the following properties*:

(1) $\nabla^F_{fX} Y = f\nabla^F_X Y$,

(2) $\nabla^F_X(fY) = X(f)Y + f\nabla^F_X Y$,

(3) $\big(\nabla^F_X(Z \circ F)\big)_p = \nabla_{F_*X_p} Z$,

*for all $X \in \Gamma(TN)$, $Y \in \Gamma(F^*TM)$, $f \in C^\infty(N)$, $Z \in \Gamma(TM)$, and $p \in N$.*

The map $\nabla^F$ is called the *connection along $F$ induced by* $\nabla$. This result will be established together with the following lemma. Consider again a moving frame $(A_1, \dots, A_m)$ on an open set $U \subset M$, and let $\Gamma^k_{ij}$ denote the respective Christoffel symbols of $\nabla$. Let $V \subset N$ be an open set such that $F(V) \subset U$.

**1.11 Lemma** *Given $X \in \Gamma(TN)$ and $Y \in \Gamma(F^*TM)$, if $X^i, Y^j \in C^\infty(V)$ are such that $F_*X = \sum_i X^i(A_i \circ F)$ and $Y = \sum_j Y^j(A_j \circ F)$ on $V$, then*

$$\nabla^F_X Y = \sum_{k=1}^m \Big(X(Y^k) + \sum_{i,j=1}^m X^i Y^j (\Gamma^k_{ij} \circ F)\Big)(A_k \circ F)$$

*on $V$.*

*Proof of Proposition 1.10 and Lemma 1.11*: Suppose first that a (possibly non-unique) map $\nabla^F$ as in the proposition exists. Like for $\nabla$, any bilinear map $\nabla^F$ satisfying (1) and (2) is defined locally on $N$. Thus, by (2),

$$\nabla^F_X Y = \sum_{j=1}^m \nabla^F_X(Y^j(A_j \circ F)) = \sum_{j=1}^m \big(X(Y^j)(A_j \circ F) + Y^j\, \nabla^F_X(A_j \circ F)\big)$$

on $V$. Furthermore, by (3) and (1),

$$\big(\nabla^F_X(A_j \circ F)\big)_p = \nabla_{F_* X_p} A_j = \sum_{i=1}^m X^i(p) \nabla_{(A_i \circ F)_p} A_j$$

for all $p \in V$. This gives the local representation stated in the lemma and shows that $\nabla^F$ is uniquely determined. Adopting this formula as a (local) definition of $\nabla^F$, one can then easily check that properties (1) to (3) are satisfied. $\square$

**1.12 Proposition** *Let $F\colon N \to M$, $\nabla$, and $\nabla^F$ be given as above.*

(1) *For $X, Y \in \Gamma(TN)$, define*

$$T^F(X,Y) := \nabla^F_X F_* Y - \nabla^F_Y F_* X - F_*[X,Y] \in \Gamma(F^*TM).$$

*Then $(T^F(X,Y))_p = T(F_* X_p, F_* Y_p)$ for all $p \in N$. In particular, if $\nabla$ is torsion-free, then $T^F \equiv 0$.*

(2) *If $\nabla$ is compatible with a Riemannian metric $g = \langle \cdot\,, \cdot \rangle$ on $M$, then*

$$X\langle V, W\rangle = \big\langle \nabla^F_X V, W \big\rangle + \big\langle V, \nabla^F_X W \big\rangle$$

*for all $X \in \Gamma(TN)$ and $V, W \in \Gamma(F^*TM)$.*

*Proof*: For (1), note that $T^F$ is a local operator, like $\nabla^F$. It thus suffices to consider vector fields $X, Y$ on $F^{-1}(U)$, where $U \subset M$ is the domain of a moving frame $(A_1, \dots, A_m)$. Then $F_* X = \sum_i X^i A_i \circ F$ and $F_* Y = \sum_j Y^j A_j \circ F$ for smooth functions $X^i, Y^j$ on $F^{-1}(U)$. Now

$$\nabla^F_X F_* Y = \sum_{j=1}^m X(Y^j)\, A_j \circ F + \sum_{j=1}^m Y^j\, \nabla^F_X (A_j \circ F),$$

and the second sum equals

$$\sum_{j=1}^m Y^j\, \nabla_{F_* X} A_j = \sum_{i,j=1}^m X^i Y^j\, (\nabla_{A_i} A_j) \circ F.$$

The term $\nabla^F_Y F_* X$ is computed similarly. Furthermore, one can check that

$$F_*[X,Y] = \sum_{j=1}^m \big(X(Y^j) - Y(X^j)\big) A_j \circ F + \sum_{i,j=1}^m X^i Y^j\, [A_i, A_j] \circ F.$$

From these identities, and since $T$ is a tensor field, it follows that

$$T^F(X,Y) = \sum_{i,j=1}^m X^i Y^j\, T(A_i, A_j) \circ F = T(F_* X, F_* Y).$$

For the proof of (2), note that the bilinear form

$$(V, W) \mapsto X\langle V, W\rangle - \langle \nabla^F_X V, W\rangle - \langle V, \nabla^F_X W\rangle$$

is $C^\infty(N)$-homogeneous in both arguments, as is readily checked. Therefore it suffices to take $V = A_i \circ F$ and $W = A_j \circ F$ for vector fields $A_i, A_j$ on an open set $U \subset M$ as above. Then

$$X\langle A_i \circ F, A_j \circ F\rangle = X(\langle A_i, A_j\rangle \circ F) = (F_* X)\langle A_i, A_j\rangle.$$

Now use the compatibility of $\nabla$ and the relation $\nabla_{F_* X} A_k = \nabla^F_X (A_k \circ F)$. □

If $\nabla = D$ is the Levi-Civita connection of $(M, g)$, then $\nabla^F = D^F$ is called the *Levi-Civita connection along $F$*. The superscript $F$ will most often be omitted. For $X = \frac{\partial}{\partial \varphi^i}$ or $X = \frac{d}{dt}$ (in the case that $N = I \subset \mathbb{R}$), we write $D_X Y$ as

$$\frac{D}{\partial \varphi^i} Y \quad \text{or} \quad \frac{D}{dt} Y,$$

repectively. A vector field $Y$ along a curve $c\colon I \to M$ is called *parallel* if

$$\frac{D}{dt} Y = \frac{D^c}{dt} Y \equiv 0.$$

In a chart $(\varphi, U)$, if $c\colon I \to U \subset M$ and $\varphi \circ c =: (x^1, \dots, x^m)$, then $(\varphi \circ c)^{\cdot} = \sum_i \dot{x}^i e_i$ and $\dot{c} = \sum_i \dot{x}^i \frac{\partial}{\partial \varphi^i} \circ c$. Hence, for a vector field $Y = \sum_j Y^j \frac{\partial}{\partial \varphi^i} \circ c$ along $c$,

$$\frac{D}{dt} Y = \sum_{k=1}^m \left( \dot{Y}^k + \sum_{i,j=1}^m \dot{x}^i Y^j (\Gamma^k_{ij} \circ c) \right) \frac{\partial}{\partial \varphi^k} \circ c$$

by Lemma 1.11. Thus $Y$ is parallel if and only if

$$\dot{Y}^k + \sum_{i,j=1}^m \dot{x}^i Y^j (\Gamma^k_{ij} \circ c) = 0$$

for $k = 1, \dots, m$. This is a linear system of ordinary differential equations of first order for $Y^1, \dots, Y^m$.

**1.13 Proposition** *Let $c\colon [a, b] \to M$ be a $C^1$-curve. Then for every $v \in TM_{c(a)}$ there exists a unique parallel vector field $Y^c_v$ along $c$ with $Y^c_v(a) = v$.*

*Proof*: This follows from the existence and uniqueness of solutions for ordinary differential equations. □

The map $P^c\colon TM_{c(a)} \to TM_{c(b)}$ defined by $P^c(v) := Y_v^c(b)$ is called *parallel transport* along $c$. This is a linear isometry, since for all $v, w \in TM_{c(a)}$,

$$\frac{d}{dt}\langle Y_v^c, Y_w^c\rangle = \Big\langle \frac{D}{dt}Y_v^c, Y_w^c\Big\rangle + \Big\langle Y_v^c, \frac{D}{dt}Y_w^c\Big\rangle = 0$$

on $[a, b]$. In general, $P^c$ depends on $c$ and not only on the end points of $c$. The *holonomy group* $\mathrm{Hol}_p$ of $(M, g)$ at $p$ is the group of all isometries of $TM_p$ of the form $P^c$ for a piecewise $C^1$ curve $c\colon [a, b] \to M$ with $c(a) = c(b) = p$. Most often, $\mathrm{Hol}_p$ acts transitively on the unit sphere $S(1) \subset TM_p$. This statement is made precise in Berger's holonomy theorem [Be1955].

## Geodesics and the exponential map

**1.14 Definition** A curve $c\colon I \to (M, g)$ with $\frac{D}{dt}\dot c \equiv 0$ is called a *geodesic*.

In particular $|\dot c|$ is constant, thus geodesics are parametrized proportionally to arc length. If $t \mapsto c(t)$ is a geodesic and $\alpha, \beta \in \mathbb{R}$, then $s \mapsto \tilde c(s) := c(\alpha s + \beta)$ is a geodesic.

Let now $c\colon [a, b] \to M$ be any smooth curve. Consider a smooth variation $\gamma\colon (-\epsilon, \epsilon) \times [a, b] \to M$ of $c = \gamma(0, \cdot\,)$. We let $V := \gamma_* \frac{\partial}{\partial s} \in \Gamma(\gamma^* TM)$ denote the variation vector field, and we put $\gamma_s := \gamma(s, \cdot\,)$ and $V_s := V(s, \cdot\,)$.

**1.15 Theorem (first variation of arc length)** *If* $|\dot c(t)| = \lambda \neq 0$ *for all* $t \in [a, b]$*, then*

$$\frac{d}{ds}\Big|_{s=0} L(\gamma_s) = \frac{1}{\lambda}\bigg(\big\langle V_0(t), \dot c(t)\big\rangle\Big|_a^b - \int_a^b \Big\langle V_0(t), \frac{D}{dt}\dot c(t)\Big\rangle\, dt\bigg).$$

*Proof*: By the second part of Proposition 1.12, for all $t \in [a, b]$,

$$\frac{\partial}{\partial s}|\dot\gamma_s(t)| = \frac{1}{2|\dot\gamma_s(t)|}\,\frac{\partial}{\partial s}\langle \dot\gamma_s(t), \dot\gamma_s(t)\rangle = \frac{1}{|\dot\gamma_s(t)|}\,\Big\langle \frac{D}{\partial s}\dot\gamma_s(t), \dot\gamma_s(t)\Big\rangle,$$

and, by the first part, $\frac{D}{\partial s}\dot\gamma = \frac{D}{\partial s}\gamma_* \frac{\partial}{\partial t} = \frac{D}{\partial t}\gamma_* \frac{\partial}{\partial s} = \frac{D}{\partial t}V$. Hence,

$$\frac{\partial}{\partial s}\Big|_{s=0} L(\gamma_s) = \frac{1}{\lambda}\int_a^b \Big\langle \frac{D}{dt}V_0(t), \dot c(t)\Big\rangle\, dt.$$

Since the integrand equals $\frac{d}{dt}\langle V_0(t), \dot c(t)\rangle - \big\langle V_0(t), \frac{D}{dt}\dot c(t)\big\rangle$, the result follows. □

The variation $\gamma$ is called *proper* if $\gamma_s(a) = c(a)$ and $\gamma_s(b) = c(b)$ for all $s$. It follows that a $C^\infty$ curve $c\colon [a, b] \to M$ is a geodesic if and only if $|\dot c|$ is constant and $\frac{d}{ds}\big|_{s=0} L(\gamma_s) = 0$ for every proper variation $\gamma$ of $c$.

In a chart $(\varphi, U)$, if $c\colon I \to U \subset M$ and $\varphi \circ c =: (x^1, \dots, x^m)$, then

$$\frac{D}{dt}\dot{c} = \sum_{k=1}^{m}\Biggl(\ddot{x}^k + \sum_{i,j=1}^{m} \dot{x}^i \dot{x}^j (\Gamma_{ij}^k \circ c)\Biggr)\frac{\partial}{\partial \varphi^k} \circ c.$$

Thus $c$ is a geodesic if and only if

$$\ddot{x}^k + \sum_{i,j=1}^{m} \dot{x}^i \dot{x}^j (\Gamma_{ij}^k \circ c) = 0$$

for $k = 1, \dots, m$. This is a non-linear system of ordinary differential equations of order two for $x^1, \dots, x^m$, called the *geodesic equation*.

**1.16 Proposition** (1) *For every $v \in TM$ there exists a geodesic $c_v\colon (\alpha_v, \omega_v) \to M$ defined on a maximal interval $(\alpha_v, \omega_v)$ around* $0$ *such that $\dot{c}_v(0) = v$.*

(2) *The set $W := \{(v, t) : v \in TM,\ t \in (\alpha_v, \omega_v)\}$ is open in $TM \times \mathbb{R}$, and the map $W \to M$, $(v, t) \mapsto c_v(t)$, is $C^\infty$.*

If $W = TM \times \mathbb{R}$, that is, $(\alpha_v, \omega_v) = (-\infty, \infty)$ for all $v \in TM$, then $(M, g)$ is called *geodesically complete*.

*Proof*: This follows from existence, uniqueness and smooth dependence of solutions on initial conditions for ordinary differential equations. □

Note that $(v, t) \in W$ if and only if $(tv, 1) \in W$, because $c_v(t) = c_{tv}(1)$. The set

$$\Omega := \{v \in TM : (v, 1) \in W\}$$

is open in $TM$.

**1.17 Definition** The $C^\infty$ map $\exp\colon \Omega \to M$ defined by $\exp(v) = c_v(1)$ is called the *exponential map* of $M$. For $p \in M$, we put $\Omega_p := \Omega \cap TM_p$ and

$$\exp_p := \exp|_{\Omega_p} : \Omega_p \to M.$$

Note that $\exp_p(tv) = c_{tv}(1) = c_v(t)$ for all $t \in (\alpha_v, \omega_v)$.

Regarding the terminology, note that if $M = S^1$ is the unit circle in $\mathbb{C}$ with the induced metric, then $TS^1_1 = \mathbb{R}i$, and $\exp_1(ti) = \cos(t) + \sin(t)i = \mathrm{e}^{ti}$ for all $t \in \mathbb{R}$. More generally, suppose that $(G, \cdot, e)$ is a Lie group with a bi-invariant Riemannian metric $\langle \cdot, \cdot \rangle$, and $X, Y$ are two left-invariant vector fields. Then the corresponding Levi-Civita connection satisfies $D_X Y = \frac{1}{2}[X, Y]$ (exercise). Thus the integral curve $c\colon \mathbb{R} \to G$ of $X$ with $c(0) = e$ is a geodesic:

$$\frac{D}{dt}\dot{c} = \frac{D}{dt}(X \circ c) = D_{\dot{c}} X = (D_X X) \circ c = 0.$$

Therefore the Lie group exponential map $X_e \mapsto c(1)$ agrees with the Riemannian exponential map, and for any subgroup $G \subset \mathrm{GL}(n, \mathbb{C})$, the former is given by the matrix exponential function $A \mapsto \mathrm{e}^A$.

**1.18 Remark** For every $p \in M$, the differential $d(\exp_p)_0$ is the identity map on $TM_p$ (where $T(TM_p)_0$ is identified with $TM_p$): for all $v \in TM_p$,

$$d(\exp_p)_0\,(v) = \frac{d}{dt}\Big|_{t=0}\exp_p(tv) = \frac{d}{dt}\Big|_{t=0}c_v(t) = v.$$

In particular, there exists an open neighborhood $V_p \subset \Omega_p$ of 0 in $TM_p$ such that $\exp_p|_{V_p} : V_p \to \exp_p(V_p) =: U_p \subset M$ is a diffeomorphism.

For a linear isometry $H\colon (TM_p, g_p) \to (\mathbb{R}^m, \langle\,\cdot\,,\cdot\,\rangle)$, this yields so-called *normal coordinates*

$$\varphi := H \circ \big(\exp_p|_{V_p}\big)^{-1} : U_p \to H(V_p) \subset \mathbb{R}^m$$

around $p$.

**1.19 Lemma** *In normal coordinates $\varphi$ around $p \in M^m$,*

$$g_{ij}(p) = \delta_{ij}, \quad \frac{\partial g_{ij}}{\partial \varphi^k}(p) = 0, \quad \textit{and} \quad \Gamma^k_{ij}(p) = 0$$

*for all $i, j, k \in \{1, \dots, m\}$.*

*Proof*: Since $d(\exp_p)_0 = \mathrm{id}$ on $TM_p$, it follows that $\frac{\partial}{\partial\varphi^i}\big|_p = d(\varphi^{-1})_0\,(e_i) = H^{-1}(e_i)$, thus $g_{ij}(p) = \delta_{ij}$. If $v \in TM_p$, $t \in \mathbb{R}$, and $tv \in V_p$, then

$$(x^1(t), \dots, x^m(t)) := \varphi(c_v(t)) = \varphi(\exp_p(tv)) = H(tv) = t\,H(v),$$

and the geodesic equation for $t \mapsto c_v(t)$ reads

$$\sum_{i,j=1}^{m} \dot{x}^i\dot{x}^j(\Gamma^k_{ij} \circ c_v) = 0$$

$(k = 1, \dots, m)$. For $t = 0$ and $v = H^{-1}(e_i + e_j)$, this gives $\Gamma^k_{ij}(p) + \Gamma^k_{ji}(p) = 0$. Since $D$ is torsion-free, $\Gamma^k_{ij} = \Gamma^k_{ji}$, so $\Gamma^k_{ij}(p) = 0$. Furthermore,

$$\frac{\partial g_{ij}}{\partial \varphi^k} = \sum_{l=1}^{m}\big(\Gamma^l_{ki}\,g_{lj} + g_{il}\Gamma^l_{kj}\big)$$

because $D$ is compatible with $g$, hence also $\frac{\partial g_{ij}}{\partial\varphi^k}(p) = 0$. □

**1.20 Proposition (Gauss lemma)** *Suppose that $v \in TM_p$ and $q := \exp_p(v)$ is defined. The differential*

$$T_v := d(\exp_p)_v : T(TM_p)_v = TM_p \to TM_q$$

*has the property that $\langle T_v(v), T_v(w)\rangle_q = \langle v, w\rangle_p$ for all $w \in TM_p$.*

Note that $T_v(v) = \frac{d}{dt}\big|_{t=1} \exp_p(tv) = \dot{c}_v(1)$.

*Proof*: Suppose that $v \neq 0$. Given $w \in TM_p$, choose $\epsilon > 0$ such that $\exp_p(v + sw)$ is defined for all $s \in (-\epsilon, \epsilon)$. Consider the family of radial geodesics $\gamma_s : [0,1] \to M$, $\gamma_s(t) := \exp_p(t(v + sw))$, where $\gamma_0 = c_v|_{[0,1]}$. By the first variation formula, Theorem 1.15,

$$\frac{d}{ds}\Big|_{s=0} L(\gamma_s) = \frac{1}{|v|}\Big\langle \frac{d}{ds}\Big|_{s=0} \gamma_s(1), \dot{c}_v(1)\Big\rangle = \frac{1}{|v|}\langle T_v(w), T_v(v)\rangle.$$

Since $L(\gamma_s) = |v + sw|$ and

$$\frac{d}{ds}\Big|_{s=0} |v + sw| = \frac{1}{|v|}\langle w, v\rangle,$$

this gives the result. □

**1.21 Proposition** *Let $p \in M$, and suppose that $r > 0$ is such that $\exp_p$ is defined on $B_r := \{v \in TM_p : |v| < r\}$ and $\exp_p|_{B_r}$ is a diffeomorphism onto $\exp_p(B_r)$.*

(1) *If $v \in B_r$ and $\gamma : [a,b] \to M$ is a piecewise $C^1$ curve from $p$ to $\exp_p(v) = c_v(1)$, then $L(\gamma) \geq L(c_v|_{[0,1]}) = |v|$, and equality holds only if $\gamma$ is a reparametrization of $c_v|_{[0,1]}$. In particular $d(p, c_v(1)) = |v|$.*

(2) *If $\varrho \in (0,r)$, and $q \in M$ is a point with $d(p,q) \geq \varrho$, then there is a $v \in B_r$ with $|v| = \varrho$ such that $\varrho + d(c_v(1), q) = d(p,q)$.*

Note that $\exp_p(B_r) \subset B(p,r) := \{q \in M : d(p,q) < r\}$ whenever $\exp_p$ is defined on $B_r \subset TM_p$. The second part of the proposition shows in particular that if $\exp_p|_{B_r}$ is a diffeomorphism, then for every $q$ with $d(p,q) = \varrho < r$ there is a $v \in B_r$ with $c_v(1) = q$, thus $\exp_p(B_r) = B(p,r)$.

*Proof*: Let $v \in B_r$ and $\gamma : [a,b] \to M$ be given as in (1). There exists a maximal $s \in (a,b]$ such that $\gamma([a,s]) \subset \exp_p(\overline{B}_{|v|})$ (compare the proof of Theorem 1.4), and the piecewise $C^1$ curve $\beta : [a,s] \to \overline{B}_{|v|}$ defined by $\beta(t) := \exp_p^{-1}(\gamma(t))$ for all $t \in [a,s]$ connects 0 to a boundary point of $\overline{B}_{|v|} \subset TM_p$. There is no loss of generality in assuming that $\beta(t) \neq 0$ for $t > a$. Then, for all but finitely many $t$, the derivative $\beta'(t)$ exists and can be written as $\beta'(t) = \lambda(t)\beta(t) + w(t)$ for some $\lambda(t) \in \mathbb{R}$ and a vector $w(t) \in TM_p$ orthogonal to $\beta(t)$. By Proposition 1.20, for $T_{\beta(t)} := d(\exp_p)_{\beta(t)}$,

$$|\gamma'(t)|^2 = \big|T_{\beta(t)}(\beta'(t))\big|^2 = |\lambda(t)\beta(t)|^2 + \big|T_{\beta(t)}(w(t))\big|^2.$$

It follows that

$$L(\gamma) \geq L(\gamma|_{[a,s]}) \geq \int_a^s |\lambda(t)\beta(t)|\, dt.$$

On the other hand,

$$L(c_v|_{[0,1]}) = |v| = |\beta(s)| = \int_a^s |\beta(t)|' \, dt,$$

where $|\beta(t)|' = \frac{1}{|\beta(t)|}\langle \beta'(t), \beta(t)\rangle = \lambda(t)|\beta(t)| \le |\lambda(t)\beta(t)|$. We conclude that $L(\gamma) \ge L(c_v|_{[0,1]})$. If equality holds, then $s = b$, and $w(t) = 0$ and $\lambda(t) \ge 0$ for all but finitely many $t$; thus $\gamma$ is a (piecewise $C^1$) reparametrization of $c_v|_{[0,1]}$.

We prove (2). Suppose that $\varrho \le d(p,q) < \infty$. Choose $v$ with $|v| = \varrho$ such that $q' := \exp_p(v)$ minimizes the distance to $q$. Let $\epsilon > 0$, and let $\gamma\colon [a,b] \to M$ be a piecewise $C^1$ curve from $p$ to $q$ of length $L(\gamma) < d(p,q) + \epsilon$. Since $q \notin \exp_p(B_\varrho)$, it follows as above that there is an $s \in [a,b]$ such that $\gamma(s) = \exp_p(v')$ for some $v'$ with $|v'| = \varrho$. By (1), $d(p,\gamma(s)) = \varrho = d(p,q')$, hence

$$d(p,q) + \epsilon > L(\gamma) \ge d(p,\gamma(s)) + d(\gamma(s),q) \ge d(p,q') + d(q',q)$$

by the choice of $q'$. Since $\epsilon > 0$ was arbitrary, it follows by the triangle inequality that $d(p,q) = d(p,q') + d(q',q) = \varrho + d(c_v(1),q)$. □

**1.22 Corollary** *Every geodesic $c\colon I \to M$ is locally length minimizing, that is, for all $t \in I$ there exists $\delta > 0$ such that if $t', t'' \in I \cap [t-\delta, t+\delta]$ and $t' < t''$, then $L(c|_{[t',t'']}) = d(c(t'), c(t''))$.*

*Proof*: Suppose that $c$ is parametrized by arc-length, and put $p := c(t)$. Let again $\Omega \subset TM$ denote the domain of exp. Note that $(\pi|_\Omega, \exp)\colon \Omega \to M \times M$ is regular at $0_p \in TM_p$, because the projection $\pi\colon TM \to M$ maps $0_q$ to $q$ for all $q$, and by Remark 1.18. It follows that there is an $\epsilon > 0$ such that $\exp_q|_{B_\epsilon}\colon B_\epsilon \to \exp_q(B_\epsilon)$ is a diffeomorphism for every $q \in B(p,\epsilon)$. Now, for any $\delta < \epsilon/2$, the result follows from the first part of Proposition 1.21. □

## The Hopf–Rinow Theorem

The following fundamental result was established in [HoR1931] for surfaces, and the generalization to higher dimensions is immediate (compare [My1935]).

**1.23 Theorem (Hopf–Rinow 1931)** *Let $(M,g)$ be a connected Riemannian manifold. Then the following are equivalent*:

(1) *$(M,d)$ is complete.*

(2) *$(M,g)$ is geodesically complete, that is,* exp *is defined on $TM$.*

(3) *There exists a point $p \in M$ such that* $\exp_p$ *is defined on $TM_p$.*

(4) *$(M,d)$ is proper*; *that is, all bounded closed subsets are compact.*

*If these properties hold, then for every pair of points $p, q \in M$ there exists a geodesic from $p$ to $q$ of length $d(p, q)$.*

A Riemannian manifold $(M, g)$ is called *complete* if $M$ is connected and properties (1) to (4) hold. Notice that every compact connected Riemannian manifold is complete.

*Proof*: Part I. Suppose first that (3) holds. We show that for every $q \in M$ there is a geodesic from $p$ to $q$ of length $d(p, q)$. Let $r > 0$ be such that $\exp_p|_{B_r}$ is a diffeomorphism onto $\exp_p(B_r)$, and suppose that $\varrho \in (0, r)$ and $d(p, q) > \varrho$. By Proposition 1.21, there exists a geodesic of length $\varrho$ from $p$ to a point $q'$ such that

$$\varrho + d(q', q) = d(p, q).$$

Then, since $\exp_p$ is defined on $TM_p$, there is a geodesic ray $c\colon [0, \infty) \to M$ emanating from $p$ such that $|\dot{c}| \equiv 1$ and $c(\varrho) = q'$. Put

$$A := \bigl\{t \in [0, d(p, q)] : t + d(c(t), q) = d(p, q)\bigr\},$$

and note that $\varrho \in A$. Let $s \in A$ be such that $\varrho \leq s < d(p, q)$. Applying Proposition 1.21 again, but with $p' := c(s)$ in place of $p$, we infer that there exist $\varrho' > 0$ and a geodesic $\sigma\colon [s, s + \varrho'] \to M$ of length $\varrho'$ from $p'$ to a point $q''$ such that

$$\varrho' + d(q'', q) = d(p', q).$$

Since $s \in A$, we get that $s + d(p', q) = d(p, q)$ and thus

$$(s + \varrho') + d(q'', q) = d(p, q).$$

Now $c|_{[0,s]}$ followed by $\sigma$ is a piecewise geodesic of length $s + \varrho' = d(p, q) - d(q'', q) \leq d(p, q'')$ from $p$ to $q''$, hence a minimizing geodesic. By the uniqueness of geodesics, $c(s + \varrho') = q''$, and so $s + \varrho' \in A$. Thus, we have shown that for every $s \in A$ with $\varrho \leq s < d(p, q)$ there is a $t \in A$ bigger than $s$. Since $A$ is closed, it follows that $d(p, q) \in A$. Hence, $c(d(p, q)) = q$.

Part II. Suppose again that (3) holds. Then it follows from Part I that $\overline{B}(p, r) \subset \exp_p(\overline{B}_r)$ for all $r > 0$, and the reverse inclusion is clear. Hence, since $\overline{B}_r$ is compact and $\exp_p$ is continuous, $\overline{B}(p, r)$ is compact as well. This shows that (3) implies (4), and clearly (4) implies (1).

Next, assume that (1) holds. Let $v \in TM$ be any unit vector, and let $c_v\colon [0, \omega_v) \to M$ be the maximal geodesic ray with inital vector $v$. Suppose that $\omega_v < \infty$. Choose a sequence $0 < t_i \uparrow \omega_v$. Then the sequence $(c_v(t_i))$ is Cauchy and thus converges to a point $q \in M$ by (1). Since the domain of exp is open in $TM$, there exists an $\epsilon > 0$ such that every geodesic $\gamma$ with $\gamma(0) \in B(q, \epsilon)$ and

$|\gamma'(0)| = 1$ is defined at least on $(-\epsilon, \epsilon)$. It follows that if $i$ is sufficiently large, so that $t_i > \omega_v - \epsilon$, then

$$d(c_v(t_i), q) = \lim_{j\to\infty} d(c_v(t_i), c_v(t_j)) \le \lim_{j\to\infty} |t_i - t_j| = \omega_v - t_i < \epsilon,$$

therefore $c_v$ can be extended to $[0, t_i + \epsilon)$. Since $t_i + \epsilon > \omega_v$, this contradicts the maximality of $c_v$. This shows that $\omega_v = \infty$. Hence, (1) implies (2), and clearly (2) implies (3).

Finally, by Part I of the proof, (2) implies the last assertion of the theorem. □

## Geodesic metric spaces

We conclude this chapter with a brief discussion of a metric analogue of the Hopf–Rinow theorem due to Stefan Cohn-Vossen [Co1935].

Let $(X, d)$ be a metric space. Recall that the *length* of a curve $c\colon I \to X$ is defined by

$$L(c) := \sup \sum_{i=1}^{k} d(c(t_{i-1}), c(t_i)) \in [0, \infty],$$

where the supremum is taken over all finite sequences $t_0 \le t_1 \le \ldots \le t_k$ in $I$. The curve $c$ has *constant speed* or is *parametrized proportionally to arc length* if there exists a constant $\lambda \ge 0$ such that $L(c|_{[t,t']}) = \lambda|t - t'|$ whenever $t, t' \in I$ and $t \le t'$, and $c$ has *unit speed* or is *parametrized by arc length* if $\lambda = 1$.

A metric space $(X, d)$ is called an *inner metric space* or a *length space* if, for every pair of points $p, q \in X$,

$$d(p, q) = \inf L(c),$$

where the infimum is taken over all curves $c\colon [0, 1] \to X$ from $p$ to $q$.

**1.24 Definition** Let $(X, d)$ be a metric space. A curve $c\colon I \to X$ is called a *geodesic* if $c$ has constant speed and for all $t \in I$ there exists a $\delta > 0$ such that $L(c|_{[t',t'']}) = d(c(t'), c(t''))$ whenever $t', t'' \in I \cap [t - \delta, t + \delta]$ and $t' < t''$. The curve $c$ is called a *minimizing geodesic* if $L(c|_{[t',t'']}) = d(c(t'), c(t''))$ for all $t', t'' \in I$ with $t' < t''$. We call $(X, d)$ a *geodesic metric space* if for every pair of points $p, q \in X$ there exists a minimizing geodesic $c\colon [0, 1] \to X$ from $p$ to $q$.

Every geodesic metric space is a length space. A partial converse is given by the following metric analogue of the Hopf–Rinow theorem.

**1.25 Theorem (Cohn-Vossen 1935)** *Let $X$ be a length space. If $X$ is locally compact and complete, then $X$ is a proper geodesic metric space.*

None of the assumptions can be omitted, as the following examples show. The length space $\mathbb{R}^2 \setminus \{0\}$ (with the induced inner metric) is locally compact, but not complete. The length space obtained from a sequence of disjoint segments $[a_i, b_i]$ with $b_i - a_i = 1 + \frac{1}{i}$, $i \in \mathbb{N}$, by gluing each $a_i$ to $a_1$ and each $b_i$ to $b_1$ is complete, but not locally compact. Neither of these length spaces is a geodesic metric space.

*Proof*: Fix a base point $z \in X$, and let $I$ denote the set of all $r \geq 0$ such that the closed ball $\bar{B}(z, r)$ is compact. Clearly $I$ is an interval containing 0. To prove that $X$ is a proper metric space, it suffices to show that $I = [0, \infty)$. Let $r \in I$. Use the local compactness of $X$ to cover the compact ball $\bar{B}(z, r)$ with finitely many balls $B(x_i, \epsilon_i)$ such that the $\bar{B}(x_i, \epsilon_i)$ are compact. Then $\bigcup \bar{B}(x_i, \epsilon_i)$ is compact and contains $\bar{B}(z, r+\delta)$ for some $\delta > 0$, hence $r+\delta \in I$. This shows that $I$ is open relative to $[0, \infty)$. To prove that $I$ is also closed, suppose that $[0, R) \subset I$, $R > 0$, and let $(y_j)_{j \in \mathbb{N}}$ be a sequence in $\bar{B}(z, R)$. Choose a decreasing sequence $(\epsilon_i)_{i \in \mathbb{N}}$ converging to 0, with $\epsilon_i < R$. Since $X$ is a length space, it follows that for all $i, j$ there exists an $x_j^i \in \bar{B}(z, R - \frac{\epsilon_i}{2})$ such that $d(x_j^i, y_j) \leq \epsilon_i$. The sequence $(x_j^1)$ has a convergent subsequence $(x_{j(1,k)}^1)$. Consider the corresponding sequence $(x_{j(1,k)}^2)$ and pick a convergent subsequence $(x_{j(2,k)}^2)$. Then consider $(x_{j(2,k)}^3)$ and select a converging subsequence $(x_{j(3,k)}^3)$. Continue in this manner. Finally, put $j(k) := j(k, k)$ for $k \in \mathbb{N}$. The sequence $(x_{j(k)}^i)_{k \in \mathbb{N}}$ converges for every $i \in \mathbb{N}$. We claim that the associated sequence $(y_{j(k)})$ is Cauchy. Let $\epsilon > 0$ and choose $i$ with $\epsilon_i \leq \epsilon$. Then $d(x_{j(k)}^i, x_{j(l)}^i) \leq \epsilon$ for $k, l$ sufficiently large. It follows that

$$\begin{aligned} d(y_{j(k)}, y_{j(l)}) &\leq d(y_{j(k)}, x_{j(k)}^i) + d(x_{j(k)}^i, x_{j(l)}^i) + d(x_{j(l)}^i, y_{j(l)}) \\ &\leq \epsilon_i + \epsilon + \epsilon_i \leq 3\epsilon. \end{aligned}$$

Since $X$ is complete, $(y_{j(k)})$ converges. This shows that every sequence $(y_j)_{j \in \mathbb{N}}$ in $\bar{B}(z, R)$ has a convergent subsequence. Hence $\bar{B}(z, R)$ is compact, that is, $[0, R] \subset I$. Thus $I$ is both open and closed in $[0, \infty)$, so $I = [0, \infty)$. This proves that $X$ is proper.

To show that $X$ is geodesic, let $p, q \in X$. Since $X$ is a length space, it follows that for every $n \in \mathbb{N}$ there exists a curve $c_n \colon [0, 1] \to X$ of constant speed from $p$ to $q$ such that $L(c_n) = d(p, q) + \frac{1}{n}$. Since the image of $c_n$ lies in the compact ball $\bar{B}(p, d(p, q) + 1)$, it follows by the Arzelà–Ascoli theorem that the $c_n$ converge to a minimizing geodesic from $p$ to $q$. $\square$

# Chapter 2

# Curvature

## The curvature tensor

**2.1 Definition** Let $(M, g = \langle \cdot, \cdot \rangle)$ be a Riemannian manifold with Levi-Civita connection $D$. The map $R \colon \Gamma(TM)^3 \to \Gamma(TM)$ defined by

$$R(X,Y)W := D_X D_Y W - D_Y D_X W - D_{[X,Y]} W$$

is called the *Riemann curvature tensor* of $(M, g)$. Briefly,

$$R(X,Y) = [D_X, D_Y] - D_{[X,Y]}.$$

The *curvature tensor* of an arbitrary connection $\nabla$ on $TM$ is defined analogously.

The opposite sign convention is also common. It is straight-forward to check that $R$ defines a $(1,3)$-tensor field: for $f \in C^\infty(M)$,

$$\begin{aligned} R(fX,Y)W &= f D_X D_Y W - D_Y(f D_X W) - D_{f[X,Y]} W + D_{Y(f)X} W \\ &= f R(X,Y)W, \end{aligned}$$

hence $R(X, fY)W = -R(fY, X)W = -fR(Y,X)W = fR(X,Y)W$, and

$$\begin{aligned} R(X,Y)(fW) &= D_X\big(Y(f)W + f D_Y W\big) - D_Y\big(X(f)W + f D_X W\big) \\ &\quad - [X,Y](f)W - f D_{[X,Y]} W \\ &= f R(X,Y)W. \end{aligned}$$

In a chart $(\varphi, U)$, with $A_i := \frac{\partial}{\partial \varphi^i}$,

$$D_{A_k} D_{A_l} A_j = D_{A_k}\Bigg(\sum_{s=1}^m \Gamma_{lj}^s A_s\Bigg) = \sum_{s=1}^m \Big(A_k(\Gamma_{lj}^s) A_s + \Gamma_{lj}^s D_{A_k} A_s\Big),$$

and $D_{A_k} A_s = \sum_i \Gamma^i_{ks} A_i$. It follows that

$$\begin{aligned} R(A_k, A_l)A_j &= \sum_{i=1}^{m} R^i_{jkl} A_i, \\ R^i_{jkl} &:= A_k(\Gamma^i_{lj}) - A_l(\Gamma^i_{kj}) + \sum_{s=1}^{m} \Big( \Gamma^s_{lj}\Gamma^i_{ks} - \Gamma^s_{kj}\Gamma^i_{ls} \Big). \end{aligned}$$

Notice the (traditional) index pattern: the upper index and the *first* lower index of $R^i_{jkl}$ represent the endomorphism $R(A_k, A_l)\colon \Gamma(TU) \to \Gamma(TU)$. By putting $R_{ijkl} := \sum_{r=1}^{m} g_{ir} R^r_{jkl}$, one gets the $\varphi$ components of $R$ as a $(0,4)$-tensor field, also denoted by $R$:

$$R\colon \Gamma(TM)^4 \to C^\infty(M), \quad R(V, W, X, Y) := \langle V, R(X, Y)W \rangle,$$

thus $R(A_i, A_j, A_k, A_l) = \langle A_i, R(A_k, A_l)A_j \rangle = \big\langle A_i, \sum_r R^r_{jkl} A_r \big\rangle = R_{ijkl}$.

**2.2 Proposition (symmetry properties)** *For all vector fields $V, W, X, Y \in \Gamma(TM)$,*

(1) $R(Y, X)W = -R(X, Y)W$ (*thus* $R(V, W, Y, X) = -R(V, W, X, Y)$),

(2) $R(X, Y)W + R(Y, W)X + R(W, X)Y = 0$,

(3) $R(W, V, X, Y) = -R(V, W, X, Y)$,

(4) $R(X, Y, V, W) = R(V, W, X, Y)$.

Property (2) is called the *first Bianchi identity*.

*Proof*: (1) clearly holds for the curvature tensor of any connection.

To verify (2) and (3), it suffices to consider local coordinate vector fields $V, W, X, Y$, so that all Lie brackets vanish. Then the cyclic sum in (2) equals

$$\begin{aligned} &D_X D_Y W - D_Y D_X W \\ &\qquad + D_Y D_W X - D_W D_Y X \\ &\qquad\qquad + D_W D_X Y - D_X D_W Y, \end{aligned}$$

which is zero since $D$ is torsion-free. Next, since $D$ is compatible with $g$,

$$XY\langle V, V \rangle = 2X\langle V, D_Y V \rangle = 2\langle D_X V, D_Y V \rangle + 2\langle V, D_X D_Y V \rangle$$

and $YX\langle V, V \rangle = 2\langle D_Y V, D_X V \rangle + 2\langle V, D_Y D_X V \rangle$. Taking the difference, we get $0 = [X, Y]\langle V, V \rangle = 2\langle V, R(X, Y)V \rangle = 2R(V, V, X, Y)$, which yields (3).

Lastly, the "pair symmetry" (4) follows algebraically from properties (1)–(3). The expression

$$\begin{aligned} &R(V, W, X, Y) + R(W, V, Y, X) - R(X, Y, V, W) - R(Y, X, W, V) \\ &+ R(V, X, Y, W) + R(W, Y, X, V) - R(X, V, W, Y) - R(Y, W, V, X) \\ &+ R(V, Y, W, X) + R(W, X, V, Y) - R(X, W, Y, V) - R(Y, V, X, W) \end{aligned}$$

vanishes, because the sum of each column is zero by (2). Furthermore, by (1) and (3), the last two rows vanish as well, whereas the first row is equal to $2R(V,W,X,Y) - 2R(X,Y,V,W)$. This gives (4). $\square$

The following result supplements Proposition 1.12.

**2.3 Proposition** *Let $R$ be the curvature tensor of a connection $\nabla$ on $TM$. For another manifold $N$ and a smooth map $F\colon N \to M$, let $R^F$ denote the curvature tensor of the connection $\nabla^F$ along $F$, given by*

$$R^F(X,Y)W := \nabla^F_X\nabla^F_Y W - \nabla^F_Y\nabla^F_X W - \nabla^F_{[X,Y]}W \in \Gamma(F^*TM)$$

*for $X,Y \in \Gamma(TN)$ and $W \in \Gamma(F^*TM)$. Then*

$$(R^F(X,Y)W)_p = R(F_*X_p, F_*Y_p)W_p$$

*for all $p \in N$.*

*Proof*: Let $A_1,\ldots,A_m$ be coordinate vector fields on an open set $U \subset M$. Suppose that $X,Y$ are vector fields on $F^{-1}(U)$, and write $F_*X = \sum_i X^i\, A_i \circ F$ and $F_*Y = \sum_j Y^j\, A_j \circ F$ for smooth functions $X^i, Y^j$ on $F^{-1}(U)$. Furthermore, suppose that $W = A \circ F$ on $F^{-1}(U)$ for a vector field $A$ on $U$, and put $B_k := \nabla_{A_k}A$. Then $\nabla^F_Y(A\circ F) = \nabla_{F_*Y}A = \sum_j Y^j\, B_j \circ F$, hence

$$\nabla^F_X\nabla^F_Y(A\circ F) = \sum_{j=1}^m X(Y^j)\, B_j \circ F + \sum_{j=1}^m Y^j\, \nabla^F_X(B_j \circ F),$$

and the second sum equals

$$\sum_{j=1}^m Y^j\, \nabla_{F_*X}B_j = \sum_{i,j=1}^m X^iY^j\, (\nabla_{A_i}B_j)\circ F.$$

The term $\nabla^F_Y\nabla^F_X(A\circ F)$ is computed similarly. Furthermore, by the formula for $F_*[X,Y]$ stated in the proof of Proposition 1.12, and since $[A_i, A_j] = 0$,

$$\nabla^F_{[X,Y]}(A\circ F) = \nabla_{F_*[X,Y]}A = \sum_{j=1}^m \big(X(Y^j) - Y(X^j)\big)\, B_j \circ F.$$

From these identities, it follows that

$$\begin{aligned} R^F(X,Y)(A\circ F) &= \sum_{i,j=1}^m X^iY^j\,(\nabla_{A_i}B_j - \nabla_{A_j}B_i)\circ F \\ &= \sum_{i,j=1}^m X^iY^j\,(R(A_i,A_j)A)\circ F \\ &= R(F_*X, F_*Y)(A\circ F). \end{aligned}$$

This gives the result. $\square$

## Sectional curvature

**2.4 Definition** Let $P$ be a 2-dimensional linear subspace of $TM_p$, and let $(v, w)$ be a basis of $P$. The number

$$\sec(P) = \frac{\langle v, R(v, w)w\rangle}{|v|^2|w|^2 - \langle v, w\rangle^2} = \frac{R(v, w, v, w)}{|v|^2|w|^2 - \langle v, w\rangle^2}$$

is called the *sectional curvature* of $P$.

The *Gram determinant* $|v|^2|w|^2 - \langle v, w\rangle^2$ gives the square of the area of the parallelogram spanned by $v$ and $w$. The number $\sec(P)$ is well-defined in that it does not depend on the choice of the basis of $P$, since the quotient is invariant under the maps $(v, w) \mapsto (w, v)$, $(v, w) \mapsto (\lambda v, w)$, and $(v, w) \mapsto (v + \lambda w, w)$ for $\lambda \in \mathbb{R}$ (use Proposition 2.2).

If the metric $g$ is multiplied by a constant, $\tilde{g} = \lambda^2 g$, where $\lambda > 0$, then the connection $D$ and hence $R$, as a $(1, 3)$-tensor field, are not affected, so

$$\sec_{\tilde{g}}(P) = \frac{\lambda^2\langle v, R(v, w)w\rangle}{\lambda^4(|v|^2|w|^2 - \langle v, w\rangle^2)} = \frac{1}{\lambda^2}\sec_g(P).$$

**2.5 Remark** For $p \in M$, let $G_2(TM_p)$ denote the set of all linear 2-planes in $TM_p$. The sectional curvature function $\sec_p\colon G_2(TM_p) \to \mathbb{R}$, together with $g_p$, determines the curvature tensor at $p$ completely. In fact, $R_p(v, w, x, y)$ can be expressed as a sum of terms of the form $\pm R_p(a, b, a, b) =: \pm S_p(a, b)$. To see this, note first that

$$\begin{aligned} 3R(x, y)w = {} & R(x, w + y)(w + y) - R(x, w)w - R(x, y)y \\ & - R(y, w + x)(w + x) + R(y, w)w + R(y, x)x, \end{aligned}$$

due to properties (1) and (2) in Proposition 2.2. Now take the inner product with $v$, and use the pair symmetry (4) and polarization for each of the six terms on the right hand side; for example, $2\langle v, R(x, w)w\rangle = 2R(v, w, x, w) = S(v + x, w) - S(v, w) - S(x, w)$.

For $\dim(M) = 2$, the function $K\colon M \to \mathbb{R}$, $K(p) = \sec(TM_p)$, is the *Gauss curvature* of $M$. In a chart $(\varphi, U)$,

$$K|_U = \frac{R_{1212}}{\det(g_{ij})}.$$

**2.6 Definition** A complete Riemannian manifold $M$ with constant sectional curvature $\sec \equiv \kappa$ is called a *space form*. A space form is called *spherical*, *flat*, or *hyperbolic* depending on whether $\kappa > 0$, $\kappa = 0$, or $\kappa < 0$, respectively.

If $\sec \equiv \kappa \in \mathbb{R}$, then the curvature tensor satisfies

$$R(X,Y)W = \kappa\big(\langle Y,W\rangle X - \langle X,W\rangle Y\big)$$

(exercise). We shall see that the model spaces

$$\begin{cases} S^m \subset (\mathbb{R}^{m+1}, \langle\,\cdot\,,\cdot\,\rangle_{\mathrm{eucl}}) & (\kappa = 1), \\ (\mathbb{R}^m, \langle\,\cdot\,,\cdot\,\rangle_{\mathrm{eucl}}) & (\kappa = 0), \\ H^m \subset (\mathbb{R}^{m,1}, \langle\,\cdot\,,\cdot\,\rangle_{\mathrm{L}}) & (\kappa = -1) \end{cases}$$

are, up to scaling, the only simply connected $m$-dimensional space forms, and every $m$-dimensional space form with $\sec \equiv \kappa \in \{-1, 0, 1\}$ is a quotient of one of these (see Theorem 4.10).

## Ricci and scalar curvature

Let $V$ be an $m$-dimensional $\mathbb{R}$-vector space; for example, $TM_p$. Then $(1,1)$-tensors $T \in V_{1,1} = V \otimes V^*$ appear in two equivalent forms:

$$T \colon V^* \times V \to \mathbb{R} \quad \text{and} \quad \widetilde{T} \colon V \to V,$$

where $T(\theta, X) = \theta(\widetilde{T}(X))$ for all $\theta \in V^*$ and $X \in V$. The *contraction* $C(T) = C(\widetilde{T}) \in \mathbb{R}$ is the *trace* of $\widetilde{T}$. With respect to a basis $(e_1, \dots, e_m)$ and the dual basis $(\epsilon^1, \dots, \epsilon^m)$,

$$T = \sum_{i,j} T^i_j e_i \otimes \epsilon^j \quad \text{and} \quad \widetilde{T} e_j = \sum_i T^i_j e_i$$

(where $(e_i \otimes \epsilon^j)(\theta, X) = \theta(e_i)\epsilon^j(X)$), hence

$$C(T) = C(\widetilde{T}) = \sum_i T^i_i = \sum_i T(\epsilon^i, e_i) = \sum_i \epsilon^i(\widetilde{T}(e_i)).$$

If the $(1,1)$-tensor $T$ is *decomposable*, that is, of the form $T = Y \otimes \omega$ for a vector $Y = \sum_i Y^i e_i$ and a covector $\omega = \sum_j \omega_j \epsilon^j$, then

$$C(Y \otimes \omega) = \sum_i (Y \otimes \omega)(\epsilon^i, e_i) = \sum_i \epsilon^i(Y)\omega(e_i) = \sum_i Y^i \omega_i = \omega(Y).$$

Next, for integers $r, s \geq 1$ and a pair of indices $k \in \{1, \dots, r\}$ and $l \in \{1, \dots, s\}$, the *contraction*

$$C = C_{k,l} \colon V_{r,s} \to V_{r-1,s-1}$$

*over* $k, l$ is defined as the contraction of the $(1,1)$-tensor obtained by fixing the arguments in all other slots. For example, if $T \in V_{2,3}$, then $C_{1,3}(T) \in V_{1,2}$ is defined by

$$(C_{1,3}T)(\theta, X, Y) = C(T(\,\cdot\,, \theta, X, Y, \cdot\,)).$$

In coordinates, the components of $C_{1,3}(T)$ are given by $T^a_{bc} = \sum_i T^{ia}_{bci}$.

Finally, for $r, s \geq n$, a general ($n$-fold) contraction

$$C\colon V_{r,s} \to V_{r-n,s-n}$$

is a composition of $n$ such (simple) contractions.

For tensor fields on a manifold $M$, contractions are defined pointwise, for every $TM_p$.

**2.7 Definition** The *Ricci curvature* of a Riemannian manifold $(M, g)$ is the contraction $\mathrm{Ric} \in \Gamma(T_{0,2}M)$ of the curvature tensor $R$ such that, for $v, w \in TM_p$, $\mathrm{Ric}(v, w) = \mathrm{Ric}_p(v, w)$ is the trace of the endomorphism $x \mapsto R(x, v)w$ of $TM_p$.

With respect to a basis $(e_1, \dots, e_m)$ of $TM_p$ and the respective dual basis $(\epsilon^1, \dots, \epsilon^m)$ of $TM_p^*$,

$$\mathrm{Ric}(v, w) = \sum_i \epsilon^i(R(e_i, v)w).$$

If $(e_1, \dots, e_m)$ is orthonormal, then $\epsilon^i = \langle e_i, \cdot\, \rangle$ and thus

$$\mathrm{Ric}(v, w) = \sum_i \langle e_i, R(e_i, v)w\rangle = \sum_i R(e_i, w, e_i, v).$$

In particular, by Proposition 2.2, $\mathrm{Ric}(v, w) = \mathrm{Ric}(w, v)$, and

$$\mathrm{Ric}(e_j, e_j) = \sum_i R(e_i, e_j, e_i, e_j) = \sum_{i\colon i \neq j} \sec(\mathrm{span}\{e_i, e_j\}).$$

In a chart $(\varphi, U)$,

$$\mathrm{Ric}\,|_U = \sum_{i,j} R_{ij}\, d\varphi^i \otimes d\varphi^j, \quad R_{ij} = \sum_k R^k_{ikj} = R_{ji}.$$

To take a further trace, we have to change the type of Ric by means of the metric $g$. Define the $(1,1)$-tensor field $\mathrm{Ric}'\colon \Gamma(TM) \to \Gamma(TM)$ such that $\mathrm{Ric}(V, W) = \langle V, \mathrm{Ric}'(W)\rangle$ for all $V, W \in \Gamma(TM)$. In a chart $(\varphi, U)$,

$$\mathrm{Ric}'\Big(\frac{\partial}{d\varphi^j}\Big) = \sum_i R^i_j \frac{\partial}{\partial\varphi^i}, \quad R^i_j = \sum_k g^{ik} R_{kj}.$$

**2.8 Definition** The *scalar curvature* of a Riemannian manifold $(M, g)$ is the contraction $\mathrm{scal} = C(\mathrm{Ric}') \in C^\infty(M)$ of $\mathrm{Ric}'$, called the *metric contraction* $C^g(\mathrm{Ric})$ of Ric.

For an orthonormal basis $(e_1, \dots, e_m)$ of $TM_p$,

$$\begin{aligned}\mathrm{scal}(p) &= \sum_j \langle e_j, \mathrm{Ric}'(e_j)\rangle = \sum_j \mathrm{Ric}(e_j, e_j) \\ &= \sum_{i,j} R(e_i, e_j, e_i, e_j) = \sum_{i,j\colon i \neq j} \sec(\mathrm{span}\{e_i, e_j\}).\end{aligned}$$

In a chart $(\varphi, U)$,

$$\operatorname{scal}|_U = \sum_i R^i_i = \sum_{i,j} g^{ij} R_{ji}.$$

If for some $p \in M$ the sectional curvature $\sec(P)$ is independent of the 2-plane $P \subset TM_p$, thus $\sec(P) = \kappa_p \in \mathbb{R}$ (compare Theorem 2.13), then

$$\operatorname{Ric}_p = (m-1)\kappa_p\, g_p \quad \text{and} \quad \operatorname{scal}(p) = m(m-1)\kappa_p.$$

Next we discuss tensor derivations.

**2.9 Proposition** *Let $\nabla$ be a connection on $TM$. There exists a unique family of bilinear operators*

$$\nabla^{r,s} \colon \Gamma(TM) \times \Gamma(T_{r,s}M) \to \Gamma(T_{r,s}M) \quad \text{for } r, s \geq 0,$$

*where $\nabla^{r,s}(X,T)$ is written as $\nabla^{r,s}_X T$ or just $\nabla_X T$, such that*

(1) $\nabla^{0,0}_X f = X(f) = df(X)$,

(2) $\nabla^{1,0}_X Y = \nabla_X Y$,

(3) $X(\theta(Y)) = \big(\nabla^{0,1}_X \theta\big)(Y) + \theta(\nabla_X Y)$, *and*

(4) $\nabla^{r+r',s+s'}_X (T \otimes T') = \big(\nabla^{r,s}_X T\big) \otimes T' + T \otimes \big(\nabla^{r',s'}_X T'\big)$.

*The following further properties hold*:

(5) $\nabla_{fX} T = f\, \nabla_X T$,

(6) $\nabla_X (fT) = X(f)T + f\, \nabla_X T$, *and*

(7) $C\big(\nabla^{r,s}_X T\big) = \nabla^{r-n,s-n}_X C(T)$ *for any $n$-fold contraction $C \colon \Gamma(T_{r,s}M) \to \Gamma(T_{r-n,s-n}M)$.*

*Proof (sketch)*: Properties (1)–(3) determine $\nabla^{0,0}, \nabla^{1,0}, \nabla^{0,1}$. It is straight-forward to check that $\big(\nabla^{0,1}_X \theta\big)(fY) = f\big(\nabla^{0,1}_X \theta\big)(Y)$ for $f \in C^\infty(M)$, so indeed $\nabla^{0,1}_X \theta \in \Gamma(T_{0,1}M)$. Property (6) follows directly from (1) and (4) since $fT = f \otimes T$. Furthermore, (6) implies that for a fixed $X$, the operators $\nabla^{r,s}_X$ are locally defined (as for $\nabla_X$). Now one can use a local representation of $T$ and (1)–(4) to show that $\nabla^{r,s}_X T$ is uniquely determined, and property (5) also follows inductively from (1)–(4). For example, let $T \in \Gamma(T_{1,1}M)$. Then $T$ is (locally) a finite sum of terms of the form $Y \otimes \omega$, and

$$\nabla^{1,1}_X (Y \otimes \omega) = \big(\nabla^{1,0}_X Y\big) \otimes \omega + Y \otimes \big(\nabla^{0,1}_X \omega\big)$$

by (4). By contracting this identity, we get that

$$C\big(\nabla^{1,1}_X (Y \otimes \omega)\big) = \omega(\nabla_X Y) + (\nabla_X \omega)(Y) = X(\omega(Y)) = \nabla^{0,0}_X(\omega(Y))$$

by (3) and (1). Since $\omega(Y) = C(Y \otimes \omega)$, this shows (7) for this case. □

**Examples**

1. Suppose that $T \in \Gamma(T_{0,2}M)$ (for example, $T = g$ or $T = \mathrm{Ric}$). Then $T(X,Y) = C(X \otimes Y \otimes T)$ for a 2-fold contraction $C$, thus

$$Z(T(X,Y)) = \nabla_Z^{0,0} C(X \otimes Y \otimes T) = C\big(\nabla_Z^{2,2}(X \otimes Y \otimes T)\big)$$

by (1) and (7). Expanding $\nabla_Z^{2,2}(X \otimes Y \otimes T)$ by (4) and applying $C$ to the individual terms, we get that

$$\begin{aligned} Z(T(X,Y)) &= C\Big((\nabla_Z X) \otimes Y \otimes T + X \otimes (\nabla_Z Y) \otimes T + X \otimes Y \otimes \big(\nabla_Z^{0,2} T\big)\Big) \\ &= T(\nabla_Z X, Y) + T(X, \nabla_Z Y) + \big(\nabla_Z^{0,2} T\big)(X,Y). \end{aligned}$$

In particular, $\nabla$ is compatible with $g$ if and only if $g$ is *parallel* with respect to $\nabla$, that is, $\nabla^{0,2} g \equiv 0$.

2. Let again $D$ denote the Levi-Civita connection of $g = \langle \cdot, \cdot \rangle$. For a function $f \in C^\infty(M)$, the *gradient* $\operatorname{grad} f \in \Gamma(TM)$ and the *Hesse form* $\mathrm{Hess}(f) \in \Gamma(T_{0,2}M)$ are defined by the relations $\langle \operatorname{grad} f, Y \rangle = df(Y)$ and

$$\begin{aligned} \mathrm{Hess}(f)(X,Y) &:= \langle D_X \operatorname{grad} f, Y \rangle = X \langle \operatorname{grad} f, Y \rangle - \langle \operatorname{grad} f, D_X Y \rangle \\ &= X(df(Y)) - df(D_X Y) \end{aligned}$$

for $X, Y \in \Gamma(TM)$. By (3), this last term is equal to $\big(D_X^{0,1} df\big)(Y)$; briefly, $\mathrm{Hess}(f) = D\,df$. Notice that the term $X(df(Y)) - df(D_X Y) = X(Y(f)) - (D_X Y)(f)$ is symmetric in $X$ and $Y$ since $D$ is torsion-free, thus $\mathrm{Hess}(f)$ is symmetric.

3. If $R$ is the curvature tensor of a connection $\nabla$ on $TM$, then $R(X,Y)W = C(W \otimes X \otimes Y \otimes R)$ for a 3-fold contraction $C$, and it follows similarly as in Example 1 that

$$\begin{aligned} \nabla_Z(R(X,Y)W) &= C\big(\nabla_Z^{4,3}(W \otimes X \otimes Y \otimes R)\big) \\ &= R(X,Y)\nabla_Z W + R(\nabla_Z X, Y)W + R(X, \nabla_Z Y)W \\ &\quad + \big(\nabla_Z^{1,3} R\big)(X,Y)W. \end{aligned}$$

A similar identity holds for $Z(R(V,W,X,Y))$ and $\nabla_Z^{0,4} R$, and if $\nabla$ is compatible with $g = \langle \cdot, \cdot \rangle$, then it follows that

$$\big\langle V, \big(\nabla_Z^{1,3} R\big)(X,Y)W \big\rangle = \big(\nabla_Z^{0,4} R\big)(V,W,X,Y).$$

Furthermore, for the Levi-Civita connection $\nabla = D$, the $(0,4)$-tensor fields $D_Z^{0,4} R$ and $R$ have the same symmetry properties (exercise).

**2.10 Lemma (second Bianchi identity)** *Let $R$ be the curvature tensor of a torsion-free connection $\nabla$. Then*

$$(\nabla_Z R)(X,Y)W + (\nabla_X R)(Y,Z)W + (\nabla_Y R)(Z,X)W = 0$$

*for all $W,X,Y,Z \in \Gamma(TM)$.*

*Proof*: It suffices to prove this for local coordinate vector fields $W,X,Y,Z$. Then we can write $R(X,Y)$ as $[\nabla_X,\nabla_Y]$, and it follows from the identity in Example 3 above that

$$\begin{aligned}
&(\nabla_Z R)(X,Y)W \\
&\quad = \nabla_Z(R(X,Y)W) - R(X,Y)\nabla_Z W - R(\nabla_Z X,Y)W - R(X,\nabla_Z Y)W \\
&\quad = \nabla_Z[\nabla_X,\nabla_Y]W - [\nabla_X,\nabla_Y]\nabla_Z W - R(\nabla_Z X,Y)W + R(\nabla_Z Y,X)W \\
&\quad = \big[\nabla_Z,[\nabla_X,\nabla_Y]\big]W - R(\nabla_Z X,Y)W + R(\nabla_Y Z,X)W,
\end{aligned}$$

where in the last step we have used that $\nabla_Z Y = \nabla_Y Z$ since $\nabla$ is torsion-free. Now it is easy to see that the cyclic sum in the lemma vanishes. $\square$

**2.11 Lemma** *Let $\nabla$ be a connection that is compatible with $g = \langle\cdot,\cdot\rangle$, and let $C = C^g\colon \Gamma(T_{0,2}M) \to C^\infty(M)$ denote the metric contraction. Then*

$$C\big(\nabla^{0,2}_X T\big) = X(C(T))$$

*for all $X \in \Gamma(TM)$ and $T \in \Gamma(T_{0,2}M)$.*

*Proof*: It suffices to prove this for an individual vector $x \in TM_p$. Choose a curve $c\colon(-\epsilon,\epsilon) \to M$ with $\dot c(0) = x$. Since $\nabla$ is compatible with $g$, there exists a parallel orthonormal frame field $(E_1,\dots,E_m)$ along $c$. Then

$$\begin{aligned}
C\big(\nabla^{0,2}_x T\big) &= \sum_{i=1}^m \big(\nabla^{0,2}_x T\big)(E_i,E_i) \\
&= \sum_{i=1}^m \big(x(T(E_i,E_i)) - T(\nabla_x E_i,E_i) - T(E_i,\nabla_x E_i)\big)
\end{aligned}$$

by Example 1. Now $\nabla_x E_i = 0$ and $\sum_i x(T(E_i,E_i)) = x(C(T))$. $\square$

The *divergence* of a vector field $Y \in \Gamma(TM)$ is the contraction

$$\operatorname{div}(Y) := C(DY) \in C^\infty(M)$$

of the $(1,1)$-tensor field $D.Y\colon \Gamma(TM) \to \Gamma(TM)$. The *divergence* of a symmetric tensor field $T \in \Gamma(T_{0,2}M)$ is the metric contraction

$$\operatorname{div}(T) := C^g_{1,2}(DT) = C^g_{1,3}(DT) \in \Gamma(TM^*)$$

of the $(0,3)$-tensor field $(D.T)(\cdot,\cdot)$.

**2.12 Proposition** *For any Riemannian manifold $M$, $d\,\mathrm{scal} = 2\,\mathrm{div}(\mathrm{Ric})$.*

*Proof*: Let $p \in M$, and let $v, w, x, y, z \in TM_p$. It follows from Lemma 2.10 and the symmetries of $R$ (see Example 3 above) that

$$\langle v, (D_z R)(x, y)w\rangle - \langle w, (D_x R)(y, z)v\rangle - \langle v, (D_y R)(x, z)w\rangle = 0.$$

Let $(e_1, \ldots, e_m)$ be an orthonormal basis of $TM_p$. By putting $(v, w, x, y) := (e_i, e_j, e_i, e_j)$ and summing over $i, j$, we get that

$$\sum_{i,j=1}^{m} \langle e_i, (D_z R)(e_i, e_j)e_j\rangle = 2 \sum_{i,j=1}^{m} \langle e_i, (D_{e_j} R)(e_i, z)e_j\rangle.$$

Since $D$ commutes with the contraction corresponding to the sum over $i$,

$$\sum_{j=1}^{m} (D_z \,\mathrm{Ric})(e_j, e_j) = 2 \sum_{j=1}^{m} (D_{e_j} \mathrm{Ric})(z, e_j).$$

Now it follows from Lemma 2.11 and the definition of the divergence that $d\,\mathrm{scal}(z) = z(\mathrm{scal}) = 2\,\mathrm{div}(\mathrm{Ric})(z)$. □

As a corollary we get a classical result due to Friedrich Schur.

**2.13 Theorem (Schur 1886)** *Let $(M, g)$ be a connected Riemannian manifold of dimension $m \geq 3$.*

(1) *If $\mathrm{Ric} = fg$ for some $f \in C^\infty(M)$, then $f$ is constant.*

(2) *If for every $p \in M$ the sectional curvature $\mathrm{sec}(P)$ is independent of $P \subset TM_p$, then $\mathrm{sec}$ is constant on $M$.*

A Riemannian manifold $(M, g)$ with $\mathrm{Ric} = \kappa g$ for some constant $\kappa \in \mathbb{R}$ is called an *Einstein manifold*. In the case $\kappa = 0$, the manifold $(M, g)$ is called *Ricci-flat*.

Note that $\mathrm{Ric} = fg$ implies that $\mathrm{scal} = mf$.

*Proof*: By Proposition 2.12,

$$m\, df = d\,\mathrm{scal} = 2\,\mathrm{div}(\mathrm{Ric}) = 2\,\mathrm{div}(fg),$$

and it is easy to check that $\mathrm{div}(fg) = df$. Since $m \geq 3$, this yields $df = 0$. As $M$ is connected, (1) follows.

For (2), note that if $\mathrm{sec}(P) = \kappa_p \in \mathbb{R}$ for all planes $P$ in $TM_p$, then $\mathrm{Ric} = fg$ for $f(p) = (m-1)\kappa_p$. Thus (2) follows from (1). □

## Curvature of submanifolds

Let $(M, g) \subset (\bar{M}, \bar{g})$ be a Riemannian submanifold with $m = \dim(M) < \bar{m} = \dim(\bar{M})$, where $g$ is the induced Riemannian metric on $M$. Thus $g_p = \bar{g}_p|_{TM_p \times TM_p}$ for every $p \in M$, where $TM_p$ is viewed as an $m$-dimensional linear subspace of $T\bar{M}_p$ in a canonical way. We let $\bar{D}$ denote the Levi-Civita connection of $\bar{M}$. The Levi-Civita connection $D$ of $M$ is then given by $D_X Y = (\bar{D}_X Y)^{\mathrm{T}}$ for all $X, Y \in \Gamma(TM)$. Furthermore, $TM^{\perp}$ denotes the normal bundle of $M$ in $\bar{M}$. For every $p$, the tangent space $T\bar{M}_p$ is the orthogonal direct sum of $TM_p$ and $TM_p^{\perp}$.

**2.14 Definition** The $\mathbb{R}$-bilinear map $h \colon \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM^{\perp})$ defined by

$$h(X, Y) = \big(\bar{D}_X Y\big)^{\perp} = \bar{D}_X Y - D_X Y$$

for all $X, Y \in \Gamma(TM)$ is called the *second fundamental form* of $M$ in $\bar{M}$.

Note that $h$ is symmetric and $C^{\infty}(M)$-homogeneous in both arguments:

$$h(Y, X) = \big(\bar{D}_Y X\big)^{\perp} = \big(\bar{D}_X Y\big)^{\perp} - [X, Y]^{\perp} = h(X, Y)$$

since $\bar{D}$ is torsion-free and $[X, Y] \in \Gamma(TM)$, and

$$h(fX, Y) = \big(\bar{D}_{fX} Y\big)^{\perp} = \big(f\bar{D}_X Y\big)^{\perp} = f\, h(X, Y).$$

If $N \in \Gamma(TM^{\perp})$ is a given unit normal field, then

$$h_N(X, Y) := \bar{g}(N, h(X, Y))$$

defines the real valued *second fundamental form of $M$ with respect to $N$*. Note that $h_N \in \Gamma(T_{0,2}M)$. The corresponding $(1, 1)$-tensor field

$$S_N : \Gamma(TM) \to \Gamma(TM), \quad g(S_N(X), Y) = h_N(X, Y),$$

is called the *shape operator* of $M$ with respect to $N$. For $p \in M$, the endomorphism of $TM_p$ determined by $S_N$ will be denoted by the same symbol $S_N$ or by $S_{N(p)}$; indeed it depends only on ($h$ and) the normal vector $N(p)$.

**2.15 Lemma** *The operator $S_{N(p)} \colon TM_p \to TM_p$ is self-adjoint. For $x \in TM_p$,*

$$S_N(x) = -(\bar{D}_x N)^{\mathrm{T}},$$

*and if the codimension $\bar{m} - m$ is one, then $S_N(x) = -\bar{D}_x N$.*

*Proof*: Since $h$ is symmetric, $S_{N(p)}$ is self-adjoint. For $X, Y \in \Gamma(TM)$,

$$g(S_N(X), Y) = \bar{g}(N, h(X, Y)) = \bar{g}(N, \bar{D}_X Y) = \bar{g}(-\bar{D}_X N, Y),$$

and in the case of codimension one, $\bar{D}_X N$ is tangent to $M$ because $2\,\bar{g}(\bar{D}_X N, N) = X\,\bar{g}(N, N) = 0$. $\square$

All these notions, $h$, $h_N$, and $S_N$, can also be defined for isometric immersions $F\colon (M,g) \to (\bar{M},\bar{g})$. For this, one replaces $\bar{D}$ by the connection $\bar{D}^F$ along $F$ and $N$ by a unit normal vector field along $F$.

**2.16 Theorem (Gauss equations)** *Let $(M,g) \subset (\bar{M},\bar{g})$ be a Riemannian submanifold as above, and let $R$ and $\bar{R}$ denote the Riemannian curvature tensors of $M$ and $\bar{M}$, respectively. Then*

$$R(V,W,X,Y) = \bar{R}(V,W,X,Y) + \bar{g}(h(V,X),h(W,Y)) - \bar{g}(h(V,Y),h(W,X))$$

*for all $V,W,X,Y \in \Gamma(TM)$.*

In particular, if $(e_1,e_2)$ is an orthonormal basis of $P \subset TM_p$, then

$$\sec_M(P) = \sec_{\bar{M}}(P) + \bar{g}(h(e_1,e_1),h(e_2,e_2)) - \bar{g}(h(e_1,e_2),h(e_1,e_2)).$$

*Proof*: Suppose that $V,W,X,Y$ are local coordinate vector fields on $M$. Then

$$R(V,W,X,Y) = g(V,R(X,Y)W) = g(V,D_XD_YW - D_YD_XW).$$

With $D_XD_YW = (\bar{D}_XD_YW)^{\mathrm{T}} = (\bar{D}_X\bar{D}_YW - \bar{D}_Xh(W,Y))^{\mathrm{T}}$ and the corresponding expression for $D_YD_XW$ we get that

$$\begin{aligned} R(V,W,X,Y) &= \bar{g}\big(V,\bar{R}(X,Y)W - \bar{D}_Xh(W,Y) + \bar{D}_Yh(W,X)\big) \\ &= \bar{R}(V,W,X,Y) + \bar{g}(\bar{D}_XV,h(W,Y)) - \bar{g}(\bar{D}_YV,h(W,X)), \end{aligned}$$

where for the second step we have used that $\bar{D}$ is compatible with $\bar{g}$ and $h(\,\cdot\,,\cdot\,)$ is normal. Since $(\bar{D}_{.}V)^{\perp} = h(V,\cdot\,)$, the result follows. □

**2.17 Lemma** *For a curve $c\colon I \to M \subset \bar{M}$ and a vector field $Y \in \Gamma(c^*TM)$ along $c$,*

$$\frac{\bar{D}}{dt}Y = \frac{D}{dt}Y + h(\dot{c},Y).$$

Note that $\frac{D}{dt}Y$ is tangential to $M$, whereas $h(\dot{c},Y)$ is normal. The lemma shows in particular that

$$\frac{\bar{D}}{dt}\dot{c} = \frac{D}{dt}\dot{c} + h(\dot{c},\dot{c}),$$

thus $c$ is an $M$-geodesic if and only if $\frac{\bar{D}}{dt}\dot{c}$ is normal to $M$.

*Proof*: We assume that $c(I)$ is contained in the domain $U \subset M$ of a moving frame $(A_1,\dots,A_m)$. Then $Y = \sum_i Y^iA_i \circ c$ for some smooth functions $Y^i$ on $I$, and

$$\begin{aligned} \frac{\bar{D}}{dt}Y - \frac{D}{dt}Y &= \sum_{i=1}^m Y^i\Big(\frac{\bar{D}}{dt}(A_i \circ c) - \frac{D}{dt}(A_i \circ c)\Big) \\ &= \sum_{i=1}^m Y^i\big(\bar{D}_{\dot{c}}A_i - D_{\dot{c}}A_i\big) \end{aligned}$$

by Proposition 1.10. Since $\bar{D}_{\dot{c}}A_i - D_{\dot{c}}A_i = h(\dot{c},A_i \circ c)$, the result follows. □

**2.18 Proposition (totally geodesic submanifolds)** *For a submanifold $M \subset \bar{M}$ the following four statements are equivalent*:

(1) $h \equiv 0$;

(2) *every geodesic in $M$ is also a geodesic in $\bar{M}$*;

(3) *if $v \in TM_p$, then the $\bar{M}$-geodesic $\bar{c}_v$ lies initially in $M$*;

(4) *if $c\colon I \to M$ is a curve, then every $D$-parallel vector field $Y \in \Gamma(c^*TM)$ is also $\bar{D}$-parallel.*

A submanifold with these properties is called *totally geodesic*.

*Proof*: Lemma 2.17 shows that (1) implies (4). Taking $Y := \dot{c}$, we see that (4) implies (2). If (2) holds, and if $v \in TM_p$, then the maximal $M$-geodesic $c_v\colon (\alpha_v, \omega_v) \to M$ is also an $\bar{M}$-geodesic, so $\bar{c}_v|_{(\alpha_v,\omega_v)} = c_v$ by uniqueness, and (3) holds. Finally, (3) implies by Lemma 2.17 that $h(v, v) = 0$ for all $v \in TM_p$, which is equivalent to (1) by the symmetry of $h$. □

For example, if $S^m$ is the unit sphere in $\mathbb{R}^{m+1}$, and $L$ is a $(k + 1)$-dimensional linear subspace of $\mathbb{R}^{m+1}$, then evidently $L \cap S^m$ is a $k$-dimensional totally geodesic submanifold of $S^m$, isometric to $S^k$. In particular, for $p \in S^m$ and a plane $P \subset TM_p$, if $L$ is the 3-dimensional subspace spanned by $p$ and $P$, then it follows from Theorem 2.16 that $\sec(P)$ equals the Gauss curvature of the 2-sphere $L \cap S^m$ (which is 1, as we know).

## Riemannian products

Let $(M, g)$ and $(M', g')$ be two Riemannian manifolds with Levi-Civita connections $D, D'$ and Riemann curvature tensors $R, R'$. Consider the product manifold $\bar{M} = M \times M'$, and recall that for $\bar{p} = (p, p') \in \bar{M}$, the tangent space splits as $T\bar{M}_{\bar{p}} = TM_p \times TM'_{p'}$.

**2.19 Definition** The *product metric* $\bar{g} := g \times g'$ on $\bar{M}$ is defined by

$$\bar{g}_{\bar{p}}(\bar{v}, \bar{w}) = g_p(v, w) + g'_{p'}(v', w')$$

for all $\bar{p} = (p, p') \in \bar{M}$ and $\bar{v} = (v, v')$, $\bar{w} = (w, w') \in T\bar{M}_{\bar{p}}$.

Notice that if $\pi\colon \bar{M} \to M$ and $\pi'\colon \bar{M} \to M'$ denote the canonical projections, then $\bar{g} = \pi^*g + (\pi')^*g'$.

It follows readily from the Koszul formula that the Levi-Civita connection $\bar{D}$ of $(\bar{M}, \bar{g})$ satisfies

$$\bar{D}_{\bar{X}}\bar{Y} = (D_X Y, D'_{X'} Y')$$

for all vector fields $\bar{X} = (X, X')$ and $\bar{Y} = (Y, Y')$ on $\bar{M}$. In particular, a curve $\bar{c} = (c, c')$ is geodesic in $\bar{M}$ if and only if $c$ is a geodesic in $M$ and $c'$ is a geodesic in $M'$. Thus $M \times \{p'\}$ and $\{p\} \times M'$ are totally geodesic submanifolds of $\bar{M}$ for all $p' \in M$ and $p \in M$.

The Riemann curvature tensor $\bar{R}$ of $(\bar{M}, \bar{g})$ is of the form

$$\bar{R}(\bar{X}, \bar{Y})\bar{W} = \big(R(X, Y)W, R'(X', Y')W'\big)$$

for $\bar{X} = (X, X')$, $\bar{Y} = (Y, Y')$, and $\bar{W} = (W, W')$. We let $\sec_M, \sec_{M'}$ and $\sec_{\bar{M}}$ denote the sectional curvature of $M, M'$ and $\bar{M}$, respectively.

**2.20 Proposition** (1) *If* $\sec_M, \sec_{M'} \geq \kappa \in \mathbb{R}$, *then* $\sec_{\bar{M}} \geq \min\{\kappa, 0\}$.

(2) *If* $\sec_M, \sec_{M'} \leq \kappa \in \mathbb{R}$, *then* $\sec_{\bar{M}} \leq \max\{\kappa, 0\}$.

(3) *For* $P = \mathrm{span}\{(v, 0), (0, w')\}$, $\sec_{\bar{M}}(P) = 0$.

*Proof*: We prove (1). Let $(\bar{v}, \bar{w}) = ((v, v'), (w, w'))$ be an orthonormal basis of the plane $P \subset T\bar{M}_{\bar{p}}$, and put

$$Q := g(v, v)\, g(w, w) - g(v, w)^2, \quad Q' := g'(v', v')\, g'(w', w') - g'(v', w')^2.$$

Now if $\sec_M, \sec_{M'} \geq \kappa$, then

$$\begin{aligned} \sec_{\bar{M}}(P) &= \bar{R}(\bar{v}, \bar{w}, \bar{v}, \bar{w}) = R(v, w, v, w) + R'(v', w', v', w') \\ &\geq (Q + Q')\kappa. \end{aligned}$$

Since $g(v, v) + g'(v', v') = \bar{g}(\bar{v}, \bar{v}) = 1$ and $g(w, w) + g'(w', w') = 1$, it follows that

$$Q + Q' \leq g(v, v)g(w, w) + (1 - g(v, v))(1 - g(w, w)) \leq 1.$$

Hence, if $\kappa < 0$, then $\sec_{\bar{M}}(P) \geq (Q + Q')\kappa \geq \kappa$, and if $\kappa \geq 0$, then $\sec_{\bar{M}}(P) \geq 0$. This shows (1), and the proof of (2) is analogous.

For (3), note that $\bar{g}((v, 0), (0, w')) = 0$, and if $((v, 0), (0, w'))$ is orthonormal, then $\sec_{\bar{M}}(P) = R(v, 0, v, 0) + R'(0, w', 0, w') = 0$. □

For example, for the sphere $(S^2, g)$ of constant curvature 1, the product $(S^2 \times S^2, g \times g)$ has sectional curvature in $[0, 1]$. Heinz Hopf asked whether or not $S^2 \times S^2$ carries a Riemannian metric of positive sectional curvature (compare p. 265 in [GrKM1975]). This is still unsolved.

# Chapter 3

# Jacobi Fields

## Second variation of arc length

Let $c\colon [a,b] \to M$ be a unit speed geodesic. For some $\epsilon > 0$, let

$$\gamma\colon (-\epsilon, \epsilon) \times [a,b] \to M, \quad \gamma_s(t) := \gamma(s,t),$$

be a *piecewise smooth variation* of $c = \gamma_0$, that is, $\gamma$ is continuous and there exists a finite subdivision $a = t_0 < t_1 < \ldots < t_k = b$ of $[a,b]$ such that $\gamma|_{(-\epsilon,\epsilon)\times[t_{i-1},t_i]}$ is $C^\infty$ for $i = 1, \ldots, k$. Put $V := \gamma_* \frac{\partial}{\partial s}$ and $V_0 := V(0,\cdot)$. We will use $V_0'$ as a short-hand for the covariant derivative $\frac{D}{dt}V_0$.

**3.1 Theorem (second variation of arc length)** *With this notation,*

$$\frac{d^2}{ds^2}\Big|_{s=0} L(\gamma_s) = \int_a^b |(V_0')^\perp|^2 - R(V_0, c', V_0, c')\,dt + \left\langle \frac{D}{\partial s}\Big|_{s=0} V, c' \right\rangle\Big|_a^b.$$

Here $(V_0')^\perp$ denotes the part of $V_0'$ normal to $c$, thus

$$|(V_0')^\perp|^2 = |V_0'|^2 - \langle V_0', c'\rangle^2.$$

Note that if the variation is *proper*, that is, if $\gamma_s(a) = c(a)$ and $\gamma_s(b) = c(b)$ for all $s$, then $\frac{D}{\partial s}\big|_{s=0} V$ vanishes for $t = a$ and $t = b$, hence

$$\frac{d^2}{ds^2}\Big|_{s=0} L(\gamma_s) = \int_a^b |(V_0')^\perp|^2 - R(V_0, c', V_0, c')\,dt.$$

If, in addition, the variation is *normal*, that is, $\langle V_0, c'\rangle = 0$, then $V_0'$ is normal, and the formula may be rewritten as

$$\frac{d^2}{ds^2}\Big|_{s=0} L(\gamma_s) = -\int_a^b \langle V_0, V_0'' + R(V_0, c')c'\rangle\,dt,$$

because $\int_a^b |V_0'|^2\,dt = \langle V_0, V_0'\rangle\big|_a^b - \int_a^b \langle V_0, V_0''\rangle\,dt$ and $V_0(a) = V_0(b) = 0$.

*Proof*: We put $T := \gamma_* \frac{\partial}{\partial t}$, thus $T(s,t) = \gamma_s'(t)$. Now $|\gamma_s'(t)| = |T(s,t)| = \langle T(s,t), T(s,t)\rangle^{1/2}$, hence

$$\frac{d}{ds}L(\gamma_s) = \int_a^b \frac{1}{|T|}\Big\langle \frac{D}{\partial s}T, T\Big\rangle\, dt,$$
$$\frac{d^2}{ds^2}L(\gamma_s) = \int_a^b \frac{1}{|T|}\Big(\Big\langle \frac{D}{\partial s}\frac{D}{\partial s}T, T\Big\rangle + \Big|\frac{D}{\partial s}T\Big|^2\Big) - \frac{1}{|T|^3}\Big\langle \frac{D}{\partial s}T, T\Big\rangle^2\, dt.$$

By the first part of Proposition 1.12, $\frac{D}{\partial s}T = \frac{D}{\partial t}V$, thus

$$\frac{d^2}{ds^2}L(\gamma_s) = \int_a^b \frac{1}{|T|}\Big(\Big|\frac{D}{\partial t}V\Big|^2 - \Big\langle \frac{D}{\partial t}V, \frac{1}{|T|}T\Big\rangle^2 + \Big\langle \frac{D}{\partial s}\frac{D}{\partial t}V, T\Big\rangle\Big)\, dt.$$

Furthermore, by Proposition 2.3,

$$\begin{aligned}\Big\langle \frac{D}{\partial s}\frac{D}{\partial t}V, T\Big\rangle &= \Big\langle \frac{D}{\partial t}\frac{D}{\partial s}V, T\Big\rangle - \langle R(T,V)V, T\rangle \\ &= \frac{\partial}{\partial t}\Big\langle \frac{D}{\partial s}V, T\Big\rangle - \Big\langle \frac{D}{\partial s}V, \frac{D}{\partial t}T\Big\rangle - R(V,T,V,T).\end{aligned}$$

Now the result follows since for $s = 0$, $|T| = |c'| = 1$ and $\frac{D}{\partial t}T = \frac{D}{dt}c' = 0$. □

As a first application we prove the following result from [Sy1936].

**3.2 Theorem (Synge 1936)** *Let $M$ be a compact connected Riemannian manifold with even dimension and positive sectional curvature. If $M$ is orientable, then $M$ is simply connected.*

*Proof*: Suppose, to the contrary, that $M$ is not simply connected. Then there exists a closed curve $\alpha\colon [0,1] \to M$ that is not homotopic to a constant curve. Since $M$ is compact, it can be shown that there exists a shortest closed unit speed curve $c\colon [0,l] \to M$ in the free homotopy class $[\alpha]$ of $\alpha$, thus $l > 0$, and $c$ is a closed geodesic (exercise). Let $H \subset TM_{c(0)}$ be the hyperplane orthogonal to $c'(0)$. Since $M$ is orientable, parallel transport along $c$ gives an orientation preserving isometry $P\colon H \to H$, and since the dimension of $H$ is odd, it follows that there exists a unit vector $v_0 \in H$ with $P(v_0) = v_0$. Now let $V_0$ be the parallel unit normal field along $c$ with $V_0(0) = v_0$ (and hence $V_0(l) = v_0$). For a variation $\gamma\colon (-\epsilon,\epsilon) \times [0,l] \to M$ of $c = \gamma_0$ with variation vector field $V_0$ along $c$,

$$\frac{d^2}{ds^2}\Big|_{s=0} L(\gamma_s) = -\int_0^l R(V_0, c', V_0, c')\, dt.$$

Since $\sec > 0$ on $M$, the integrand is positive, thus $L(\gamma_s)$ attains a strict local maximum at $s = 0$, in contradiction to $c$ being a shortest curve in $[\alpha]$. □

**3.3 Remark** Every non-orientable manifold $M$ has a two-sheeted *orientable covering* $\pi\colon \bar{M} \to M$. Hence, if $M$ is as in the above theorem, but non-orientable, then one can consider $(\bar{M}, \bar{g})$ instead of $(M, g)$, where $\bar{g} := \pi^* g$. It follows that $\bar{M}$ is simply connected, thus $\pi$ is the universal covering, and the fundamental group of $M$ has order two. As a consequence, for any compact manifold $N$ with non-trivial fundamental group, for example $\mathbb{R}P^n$, $N \times N$ cannot carry a Riemannian metric with positive sectional curvature, because $\pi_1(N \times N) = \pi_1(N) \times \pi_1(N)$ has order at least four.

## Jacobi fields

**3.4 Definition** A vector field $Y$ along a geodesic $c\colon I \to M$ (where, as usual, $I$ is any interval with non-empty interior) is called a *Jacobi field* if it satisfies the *Jacobi equation*

$$\frac{D}{dt}\frac{D}{dt}Y + R(Y, c')c' = 0;$$

in brief, $Y'' + R(Y, c')c' = 0$.

Notice that if $t \mapsto Y(t)$ is a Jacobi field along $t \mapsto c(t)$, and if $\alpha, \beta \in \mathbb{R}$, then $s \mapsto \tilde{Y}(s) := Y(\alpha s + \beta)$ is a Jacobi field along $s \mapsto \tilde{c}(s) := c(\alpha s + \beta)$.

**3.5 Lemma** *The set of Jacobi fields along a geodesic $c\colon I \to M^m$ is a $2m$-dimensional vector space. For $t_0 \in I$ and $v, w \in TM_{c(t_0)}$ there is a unique Jacobi field $Y$ along $c$ with $Y(t_0) = v$ and $Y'(t_0) = w$.*

*Proof*: With respect to a parallel orthonormal frame $(E_1, \ldots, E_m)$ along $c$, the Jacobi equation corresponds to a system of linear ordinary differential equations of second order for the functions $Y^i := \langle Y, E_i \rangle$, with $Y = \sum_i Y^i E_i$. Indeed, $Y' = \sum_i (Y^i)' E_i$ and $Y'' = \sum_i (Y^i)'' E_i$, and $R(Y, c')c' = \sum_j Y^j R(E_j, c')c' = \sum_{i,j} Y^j \varrho^i_j E_i$ for some functions $\varrho^i_j$; thus $Y$ is a Jacobi field if and only if

$$(Y^i)'' + \sum_{j=1}^{m} \varrho^i_j Y^j = 0$$

for $i = 1, \ldots, m$. This gives the result. □

**3.6 Proposition** *Let $c\colon [0, l] \to M$ be a unit speed geodesic. If $\gamma$ is a variation of $c = \gamma_0\colon [0, l] \to M$ such that $\gamma_s = \gamma(s, \cdot)$ is a geodesic for every $s$, then the variation vector field $Y := V_0$ is a Jacobi field. Conversely, every Jacobi field along $c$ is of this form.*

*Proof*: For the first part, put again $V := \gamma_* \frac{\partial}{\partial s}$ and $T := \gamma_* \frac{\partial}{\partial t}$, as in the proof of Theorem 3.1. By Propositions 1.12 and 2.3,

$$\frac{D}{\partial t}\frac{D}{\partial t}V = \frac{D}{\partial t}\frac{D}{\partial s}T = \frac{D}{\partial s}\frac{D}{\partial t}T - R(V,T)T.$$

Since all curves $\gamma_s$ are geodesics, $\frac{D}{\partial t}T = 0$, thus $V$ satisfies the Jacobi equation.

Conversely, suppose that $Y$ is a Jacobi field along $c$. Choose a curve $\sigma\colon(-\epsilon,\epsilon) \to M$ with $\sigma(0) = c(0)$ and $\sigma'(0) = Y(0)$. Let $X$ and $W$ be any vector fields along $\sigma$ such that $X(0) = c'(0)$ and $W(0) = Y'(0) - \frac{D}{ds}\big|_{s=0}X$. For $s \in (-\epsilon,\epsilon)$ sufficiently close to 0, the geodesic $\gamma_s\colon [0,l] \to M$ with

$$\gamma_s(0) = \sigma(s) \quad \text{and} \quad \gamma_s'(0) = \Gamma(s) := X(s) + sW(s)$$

is defined; note that $\gamma_0 = c$. By the first part, the corresponding variation vector field $V_0$ along $c$ is a Jacobi field. In view of Lemma 3.5, we just need to show that $V_0$ and $Y$ satisfy the same initial conditions; then $V_0 = Y$ and thus $Y$ is a Jacobi field. First, $V_0(0) = \sigma'(0) = Y(0)$ by the choice of $\sigma$. Secondly,

$$V_0'(0) = \frac{D}{dt}\Big|_{t=0} V_0 = \frac{D}{ds}\Big|_{s=0}\Gamma = \frac{D}{ds}\Big|_{s=0} X + W(0) = Y'(0)$$

by the choice of $X$ and $W$. $\square$

**3.7 Remark** If $Y$ is a Jacobi field along the geodesic $c\colon I \to M$, then

$$\langle Y,c'\rangle'' = \langle Y',c'\rangle' = \langle Y'',c'\rangle = -\langle R(Y,c')c',c'\rangle = 0,$$

thus $\langle Y,c'\rangle$ is just an affine function. It follows that the tangential and normal parts

$$Y^{\mathrm{T}} = \frac{1}{|c'|^2}\langle Y,c'\rangle c', \quad Y^{\perp} = Y - Y^{\mathrm{T}}$$

are Jacobi fields as well, because $(Y^{\mathrm{T}})'' = 0$ and $R(Y^{\mathrm{T}},c')c' = 0$.

**Example** Let $c\colon \mathbb{R} \to M$ be a unit speed geodesic in a space form of curvature $\kappa \in \mathbb{R}$, and let $E \in \Gamma(c^*TM)$ be a parallel unit normal field along $c$. For a function $f\colon \mathbb{R} \to \mathbb{R}$, the normal field $Y := fE$ is a Jacobi field along $c$ if and only if

$$f'' + \kappa f = 0.$$

The solution with initial condition $(f(0), f'(0)) = (a,b) \in \mathbb{R}^2$ is $f = a\,\mathrm{cs}_\kappa + b\,\mathrm{sn}_\kappa$; see Chapter 5 for a discussion of these functions $\mathrm{cs}_\kappa, \mathrm{sn}_\kappa\colon \mathbb{R} \to \mathbb{R}$.

We now prove the following result from [My1941]. The two-dimensional case goes back to Bonnet (1855).

**3.8 Theorem (Myers 1941)** *Let $(M, g)$ be a complete Riemannian manifold with* $\mathrm{Ric}(v, v) \ge (m-1)\kappa$ *for all unit vectors $v \in TM$ and for some constant $\kappa > 0$. Then*

$$\mathrm{Diam}(M) := \sup\{d(p,q) : p, q \in M\} \le \frac{\pi}{\sqrt{\kappa}},$$

*in particular $M$ is compact and the fundamental group $\pi_1(M)$ is finite.*

Note that both bounds, for the Ricci curvature as well as the diameter, are attained if $M$ is the $m$-dimensional model space of constant sectional curvature $\kappa > 0$, the sphere of radius $\frac{1}{\sqrt{\kappa}}$. See also Theorem 6.8.

*Proof*: Let $p, q \in M$ with $l := d(p,q) > 0$. By completeness, there exists a unit speed geodesic $c \colon [0, l] \to M$ from $p$ to $q$. Choose a parallel orthornormal frame $(E_1, \dots, E_m)$ along $c$ with $E_m = c'$. Put $\lambda := \left(\frac{\pi}{l}\right)^2$, and note that the function $t \mapsto \mathrm{sn}_\lambda(t) := \frac{1}{\sqrt{\lambda}} \sin(\sqrt{\lambda} t)$ vanishes at 0 and $l = \frac{\pi}{\sqrt{\lambda}}$ and is positive on $(0, l)$. For $i = 1, \dots, m-1$, consider a variation of $c$ with variation vector field $t \mapsto \mathrm{sn}_\lambda(t)\, E_i(t)$ along $c$. Since $c$ has length $d(p,q)$, it follows from the second variation formula (Theorem 3.1) that

$$0 \le -\int_0^l \left\langle \mathrm{sn}_\lambda E_i,\ \mathrm{sn}_\lambda'' E_i + R(\mathrm{sn}_\lambda E_i, c')c' \right\rangle dt,$$

where $\mathrm{sn}_\lambda'' = -\lambda\, \mathrm{sn}_\lambda$. Hence,

$$\int_0^l \langle E_i, R(E_i, c')c' \rangle\, dt \le \lambda l.$$

By taking the sum from $i = 1$ to $m-1$ and invoking the bound on the Ricci curvature, we get that $(m-1)\kappa l \le (m-1)\lambda l$, thus $l = \frac{\pi}{\sqrt{\lambda}} \le \frac{\pi}{\sqrt{\kappa}}$. Since $p \ne q$ were arbitrary, this shows that $\mathrm{Diam}(M) \le \frac{\pi}{\sqrt{\kappa}}$.

In particular, $M$ is complete and bounded, hence compact. Furthermore, if $\pi \colon \tilde{M} \to M$ is the universal covering of $M$, then $\tilde{M}$, equipped with the Riemannian covering metric $\tilde{g} = \pi^* g$, satisfies the assumptions of the theorem as well, hence $\tilde{M}$ is compact, and thus $\pi_1(M)$ is finite. □

## Conjugate points

**3.9 Definition** Let $c \colon [a, b] \to M$ be a geodesic from $p$ to $q$ (where $a < b$, but not necessarily $p \ne q$). Then $q$ is said to be *conjugate* to $p$ along $c$ if there exists a non-trivial (that is, not everywhere vanishing) Jacobi field $Y$ along $c$ with $Y(a) = 0 = Y(b)$.

Note that this notion is independent of the choice of (constant speed) parametrization of $c$. Note also that for any such Jacobi field $Y$ with $Y(a) = 0 = Y(b)$, the affine function $\langle Y, c' \rangle$ vanishes, thus $Y$ is normal to $c$.

For example, antipodal points on the sphere $S^m(r) \subset \mathbb{R}^{m+1}$ are conjugate along any arc of a great circle connecting them.

**3.10 Remark** For $c$ as above, if $q$ is *not* conjugate to $p$ along $c$, then the linear map that sends every Jacobi field $Y$ along $c$ to the pair $(Y(a), Y(b)) \in TM_p \times TM_q$ has a trivial kernel and is thus an isomorphism of $2m$-dimensional vector spaces. Hence, for every pair $(v, w) \in TM_p \times TM_q$, there is a unique Jacobi field $Y$ along $c$ with $Y(a) = v$ and $Y(b) = w$.

**3.11 Lemma** *Let $c\colon [0, l] \to M$ be a unit speed geodesic from $p$ to $q$ with initial vector $c'(0) =: v$. Then $q$ is conjugate to $p$ along $c$ if and only if $lv$ is a singular point of* $\exp_p$.

*Proof*: The unique Jacobi field $Y$ along $c$ with initial conditions $Y(0) = 0$ and $Y'(0) = w \in TM_p$ is given by $Y(t) = d(\exp_p)_{tv}(tw)$. This is the variation vector field along $c$ of the variation defined by $\gamma_s(t) := \exp_p(t(v + sw))$, compare the proof of Proposition 3.6.

Hence, if $Y$ is a non-trivial Jacobi field along $c$ with $Y(0) = 0 = Y(l)$, then $w := Y'(0) \neq 0$ since $Y \not\equiv 0$, and thus the differential $d(\exp_p)_{lv}$ is singular since it maps $lw$ to $Y(l) = 0$. Conversely, if $w \in TM_p$ is any non-zero vector such that $d(\exp_p)_{lv}(lw) = 0$, then the Jacobi field defined by $Y(t) := d(\exp_p)_{tv}(tw)$ satisfies $Y(0) = 0 = Y(l)$ and is non-trivial since $Y'(0) = w \neq 0$. □

**3.12 Theorem** *Let $c\colon [0, l] \to M$ be a unit speed geodesic from $p$ to $q$. Suppose that no $c(t)$ with $t \in (0, l]$ is conjugate to $p$ along $c|_{[0,t]}$. Then there exists an $\epsilon > 0$ such that $L(\gamma) \geq L(c)$ for every piecewise $C^1$ curve $\gamma\colon [0, l] \to M$ from $p$ to $q$ that satisfies $d(\gamma(t), c(t)) \leq \epsilon$ for all $t$; furthermore, equality $L(\gamma) = L(c)$ holds only if $\gamma$ is a reparametrization of $c$.*

It is not true in general that $L(\gamma) \geq L(c)$ for *all* $C^1$ curves $\gamma$ from $p$ to $q$. For example, on the (flat) cylinder $S^1(r) \times \mathbb{R} \subset \mathbb{R}^3$, there are no pairs of conjugate points at all, yet there are non-constant closed geodesics, so that even $p = q$.

*Proof*: Let $v := c'(0)$, so that $c(t) = \exp_p(tv)$. By the assumption and Lemma 3.11, $d(\exp_p)_{tv}\colon TM_p \to TM_{c(t)}$ is bijective for all $t \in [0, l]$. It follows that there exist $0 = t_0 < t_1 < \ldots < t_k = l$ and open sets $U_1, \ldots, U_k \subset TM_p$ such that $tv \in U_i$ for $t \in [t_{i-1}, t_i]$ and $\exp_p|_{U_i}$ is a diffeomorphism onto the open set $V_i := \exp_p(U_i)$ for $i = 1, \ldots, k$. Choose $\epsilon > 0$ such that every $V_i$ contains the closed $\epsilon$-neighborhood of $c([t_{i-1}, t_i])$. Now let $\gamma\colon [0, l] \to M$ be a piecewise $C^1$ curve from $p$ to $q$ with $d(\gamma(t), c(t)) \leq \epsilon$ for all $t \in [0, l]$. Then $\gamma([t_{i-1}, t_i]) \subset V_i$, and we can define a curve

$\beta \colon [0,l] \to TM_p$ from 0 to $lv$ such that $\beta(t) = (\exp_p|_{U_i})^{-1}(\gamma(t))$ for $t \in [t_{i-1}, t_i]$. Note that $\beta$ is piecewise $C^1$, and $\exp_p \circ \beta = \gamma$. Now the proof can be completed as for the first part of Proposition 1.21. $\square$

**3.13 Definition** Let $c \colon [a,b] \to M$ be a unit speed geodesic. The *index form* $I = I_c$ of $c$ is the symmetric bilinear form on the space of all piecewise smooth normal vector fields along $c$ defined by

$$\begin{aligned} I(X,Y) &:= \int_a^b \langle X',Y'\rangle - R(X,c',Y,c')\,dt \\ &= \int_a^b \langle X,Y'\rangle' - \big\langle X, Y'' + R(Y,c')c'\big\rangle\,dt. \end{aligned}$$

Note that if $Y$ is the variation vector field along $c$ of a piecewise smooth, proper and normal variation $\gamma$ of $c = \gamma_0$, then

$$\frac{d^2}{ds^2}\Big|_{s=0} L(\gamma_s) = I(Y,Y).$$

**3.14 Lemma** *Let $c \colon [a,b] \to M$ be a unit speed geodesic. If $Y$ is a normal Jacobi field along $c$, then $I(X,Y) = 0$ for every piecewise smooth normal field $X$ along $c$ with $X(a) = 0 = X(b)$. Conversely, if $Y$ is a piecewise smooth normal field along $c$, and $I(X,Y) = 0$ for every smooth normal field $X$ along $c$ with $X(a) = 0 = X(b)$, then $Y$ is a (smooth) Jacobi field.*

*Proof*: By the second expression for the index form, if $Y$ is a normal Jacobi field along $c$, then $I(X,Y) = \langle X,Y'\rangle\big|_a^b = 0$ for every piecewise smooth normal field $X$ along $c$ with $X(a) = 0 = X(b)$.

Conversely, suppose that $Y$ is a piecewise smooth normal field along $c$, and $I(X,Y) = 0$ for every smooth normal field $X$ along $c$ with $X(a) = 0 = X(b)$. Let $a = t_0 < t_1 < \ldots < t_k = b$ be such that $Y_i := Y|_{[t_{i-1},t_i]}$ is smooth for $i = 1,\ldots,k$. First, choose a smooth function $f$ on $[a,b]$ that vanishes at $t_0,\ldots,t_k$ and is positive elsewhere. Then $X := f\,(Y'' + R(Y,c')c')$ is smooth and

$$0 = I(X,Y) = -\int_a^b f\,\big|Y'' + R(Y,c')c'\big|^2\,dt;$$

thus $Y_i$ is a Jacobi field for $i = 1,\ldots,k$. Secondly, choose a smooth normal vector field $W$ along $c$ such that $W(a) = 0 = W(b)$ and $W(t_i) = Y_i'(t_i) - Y_{i+1}'(t_i)$ for $i = 1,\ldots,k-1$. Then

$$0 = I(W,Y) = \sum_{i=1}^{k} \langle W, Y_i'\rangle\big|_{t_{i-1}}^{t_i} = \sum_{i=1}^{k-1} |W(t_i)|^2,$$

thus $Y_i'(t_i) = Y_{i+1}'(t_i)$ for $i = 1,\ldots,k-1$, and it follows from Lemma 3.5 that $Y$ is in fact a (smooth) Jacobi field along $c$. $\square$

The following result complements Theorem 3.12.

**3.15 Theorem** *Let $c\colon [0,l] \to M$ be a unit speed geodesic. If there exists an $r \in (0,l)$ such that $c(r)$ is conjugate to $c(0)$ along $c|_{[0,r]}$, then there exists a piecewise smooth, proper and normal variation $\gamma\colon (-\epsilon,\epsilon) \times [0,l] \to M$ of $c = \gamma_0$ such that $L(\gamma_s) < L(c)$ for all $s \in (-\epsilon,\epsilon) \setminus \{0\}$.*

*Proof*: Let $Y \not\equiv 0$ be a Jacobi field along $c|_{[0,r]}$ with $Y(0) = 0 = Y(r)$. Note that $Y$ and $Y'$ are normal. Choose a smooth normal field $X$ along $c$ with $X(0) = 0 = X(l)$ and $X(r) = -Y'(r)$, and put $Y(t) := 0$ for $t \in (r,l]$. Then $I(X,Y) = \langle X(r), Y'(r)\rangle = -|Y'(r)|^2 < 0$ and $I(Y,Y) = 0$, hence

$$I(\lambda X + Y, \lambda X + Y) = \lambda^2 I(X,X) - 2\lambda\,|Y'(r)|^2 < 0$$

for any sufficiently small $\lambda > 0$. Then $\frac{d^2}{ds^2}\big|_{s=0} L(\gamma_s) < 0$ for a variation $\gamma$ with (piecewise smooth) variation vector field $\lambda X + Y$ along $c$. $\square$

## The Rauch Comparison Theorem

**3.16 Proposition (first index lemma)** *Let $c\colon [0,l] \to M$ be a unit speed geodesic, and suppose that no $c(t)$ with $t \in (0,l]$ is conjugate to $c(0)$ along $c|_{[0,t]}$. If $X$ is a piecewise smooth normal vector field along $c$ and $Y$ is the unique Jacobi field along $c$ such that $Y(0) = X(0)$ and $Y(l) = X(l)$, then*

$$I(X,X) \geq I(Y,Y),$$

*and equality holds if and only if $X = Y$.*

Note that $Y$ exists and is normal according to Remarks 3.10 and 3.7.

*Proof*: Put $V_0 := X - Y$ and consider a (piecewise smooth, proper and normal) variation of $c$ with variation vector field $V_0$ along $c$. It follows from Theorem 3.12 that $I(V_0,V_0) \geq 0$. Furthermore, $I(X,Y) - I(Y,Y) = I(V_0,Y) = 0$ by the first part of Lemma 3.14, hence

$$I(X,X) - I(Y,Y) = I(X - Y, X - Y) = I(V_0,V_0) \geq 0.$$

Suppose now that $I(X,X) = I(Y,Y)$. Then $I(V_0,V_0) = 0$ for $V_0 = X - Y$ as above. Let $W$ be any smooth normal field along $c$ with $W(0) = 0 = W(l)$. For every $\lambda \in \mathbb{R}$,

$$2\lambda\, I(V_0,W) + \lambda^2 I(W,W) = I(V_0 + \lambda W, V_0 + \lambda W) \geq 0,$$

again by Theorem 3.12. Thus $I(V_0,W) = 0$ for all such $W$, and so $V_0$ is a Jacobi field by the second part of Lemma 3.14. Since $V_0(0) = 0 = V_0(l)$ and $c(l)$ is not conjugate to $c(0)$ along $c$, we conclude that $V_0 \equiv 0$, that is, $X = Y$. $\square$

We now state the common assumptions for the two comparison theorems of Rauch [Ra1951] and Berger [Be1962] and a corollary.

**3.17 Assumptions** Suppose that $M$ and $\bar{M}$ are Riemannian manifolds with $\dim(M) \leq \dim(\bar{M})$, and $c\colon [0,l] \to M$ and $\bar{c}\colon [0,l] \to \bar{M}$ are unit speed geodesics such that

$$\sec_M(P) \leq \sec_{\bar{M}}(\bar{P})$$

whenever $t \in [0,l]$, $c'(t) \in P$, and $\bar{c}\,'(t) \in \bar{P}$.

**3.18 Theorem (Rauch 1951)** *Suppose, in addition to Assumptions* 3.17*, that no $\bar{c}(t)$ with $t \in (0,l)$ is conjugate to $\bar{c}(0)$ along $\bar{c}|_{[0,t]}$. If $Y$ and $\bar{Y}$ are Jacobi fields along $c$ and $\bar{c}$, respectively, such that $Y(0) = 0 = \bar{Y}(0)$ and*

$$|Y'(0)| = |\bar{Y}'(0)|, \quad \langle Y'(0), c'(0)\rangle = \langle \bar{Y}'(0), \bar{c}\,'(0)\rangle,$$

*then $|Y(t)| \geq |\bar{Y}(t)|$ for all $t \in [0,l]$.*

In particular, no $c(t)$ with $t \in (0,l)$ is conjugate to $c(0)$ along $c|_{[0,t]}$.

*Proof*: We assume that $Y, \bar{Y}$ are normal along $c, \bar{c}$, as the theorem follows readily from the result in this special case by virtue of Remark 3.7. (In the general case, $|Y^{\mathrm{T}}| = |\bar{Y}^T|$ because $\langle Y, c'\rangle(0) = \langle \bar{Y}, c'\rangle(0)$ and $\langle Y, c'\rangle'(0) = \langle \bar{Y}, \bar{c}\,'\rangle'(0)$.) We further assume that $\bar{Y} \not\equiv 0$, thus $\bar{Y}'(0) \neq 0$. Note that then $\bar{Y}(t) \neq 0$ for all $t \in (0,l)$ since conjugate points are excluded.

Fix $r \in (0,l)$ for the moment, and put $\lambda := |Y(r)|$ and $\bar{\lambda} := |\bar{Y}(r)|$. Choose a parallel orthonormal frame $(E_1, \ldots, E_{m-1}, c')$ along $c$ such that $\lambda\, E_1(r) = Y(r)$. Since $\dim(\bar{M}) \geq \dim(M)$, there is an orthonormal system $(\bar{E}_1, \ldots, \bar{E}_{m-1}, \bar{c}')$ of parallel vector fields along $\bar{c}$ such that $\bar{\lambda}\, \bar{E}_1(r) = \bar{Y}(r)$. Now consider the vector field $\bar{X} := \sum_{i=1}^{m-1} \langle Y, E_i\rangle \bar{E}_i$ along $\bar{c}$. Then $|\bar{X}| = |Y|$ and $|\bar{X}'| = |Y'|$, furthermore

$$\bar{\lambda}\, \bar{X}(r) = \bar{\lambda}\lambda\, \bar{E}_1(r) = \lambda\, \bar{Y}(r).$$

Let $I^r$ and $\bar{I}^r$ denote the index forms of $c|_{[0,r]}$ and $\bar{c}|_{[0,r]}$. By the assumption involving the sectional curvatures,

$$\begin{aligned} I^r(Y,Y) &= \int_0^r |Y'|^2 - R_M(Y, c', Y, c')\, dt \\ &\geq \int_0^r |\bar{X}'|^2 - R_{\bar{M}}(\bar{X}, \bar{c}', \bar{X}, \bar{c}')\, dt = \bar{I}^r(\bar{X}, \bar{X}). \end{aligned}$$

Hence, $\bar{\lambda}^2 I^r(Y,Y) \geq \bar{\lambda}^2 \bar{I}^r(\bar{X}, \bar{X}) \geq \lambda^2 \bar{I}^r(\bar{Y}, \bar{Y})$ by Proposition 3.16. Evaluating the index forms using the second expression in Definition 3.13, we thus get that

$$|\bar{Y}(r)|^2 \langle Y(r), Y'(r)\rangle \geq |Y(r)|^2 \langle \bar{Y}(r), \bar{Y}'(r)\rangle.$$

Since $r \in (0,l)$ was arbitrary, it follows that $\big(|Y|^2/|\bar{Y}|^2\big)' \geq 0$ on $(0,l)$. Furthermore, using L'Hôpital's rule twice, one can check that $|Y|^2/|\bar{Y}|^2$ tends to 1 as $t \to 0$ (this uses that $Y(0) = 0 = \bar{Y}(0)$ and $|Y'(0)| = |\bar{Y}'(0)| \neq 0$). Thus $|Y| \geq |\bar{Y}|$ on $[0,l]$. □

**3.19 Corollary** *Let again $M, \bar{M}$ and $c, \bar{c}$ be given as in Assumptions* 3.17 *and suppose, as in Theorem* 3.18*, that no $\bar{c}(t)$ with $t \in (0, l)$ is conjugate to $\bar{c}(0)$ along $\bar{c}|_{[0,t]}$. Put $p := c(0)$, $v := c'(0)$, $\bar{p} := \bar{c}(0)$, $\bar{v} := \bar{c}\,'(0)$, and let $H\colon TM_p \to T\bar{M}_{\bar{p}}$ be a linear isometric embedding such that $H(v) = \bar{v}$. Then for all $w \in TM_p$ and $\bar{w} := H(w)$,*

$$\big|d(\exp_p)_{lv}(w)\big| \geq \big|d(\exp_{\bar{p}})_{l\bar{v}}(\bar{w})\big|.$$

*Proof*: Let $Y$ and $\bar{Y}$ be the Jacobi fields along $c$ and $\bar{c}$, respectively, with $Y(0) = 0$, $Y'(0) = w$, and $\bar{Y}(0) = 0$, $\bar{Y}'(0) = \bar{w}$. Then $|Y'(0)| = |w| = |\bar{w}| = |\bar{Y}'(0)|$ and

$$\langle Y'(0), c'(0)\rangle = \langle w, v\rangle = \langle \bar{w}, \bar{v}\rangle = \langle \bar{Y}'(0), \bar{c}\,'(0)\rangle.$$

Furthermore, $Y(t) = d(\exp_p)_{tv}(tw)$ and $\bar{Y}(t) = d(\exp_{\bar{p}})_{t\bar{v}}(t\bar{w})$ (see the proof of Lemma 3.11). Now Theorem 3.18 shows that $|Y(l)| \geq |\bar{Y}(l)|$. This gives the result. $\square$

Corollary 3.19 may further be used to compare distances or volumes; see, for example, the proofs of Theorems 5.12 and 6.1.

## Focal points and the Rauch–Berger Theorem

Let $M$ be a Riemannian manifold with Levi-Civita connection $D$, let $\Sigma \subset M$ be a submanifold, and let $N \in \Gamma(T\Sigma^\perp)$ be a unit normal field. Suppose that $c\colon [0, l] \to M$ is a unit speed geodesic from $p \in \Sigma$ to $q \in M$ with $c'(0) = N(p)$. Recall that the shape operator of $(\Sigma, N)$ at $p$ is the linear operator $S_N = S_{N(p)}\colon T\Sigma_p \to T\Sigma_p$ satisfying $S_N(x) = -(D_x N)^{\mathrm{T}}$ (Lemma 2.15).

**3.20 Proposition** *For a Jacobi field $Y$ along $c$, the following are equivalent*:

(1) *$Y$ is the variation vector field along $c$ of a variation $\gamma$ of $c = \gamma_0$ such that every curve $\gamma_s = \gamma(s, \cdot\,)$ is a geodesic with $\gamma_s(0) \in \Sigma$ and $\gamma_s'(0) \in T\Sigma^\perp_{\gamma_s(0)}$.*

(2) *$Y(0) \in T\Sigma_p$ and $Y'(0)^{\mathrm{T}} = (D_{Y(0)}N)^{\mathrm{T}} = -S_N(Y(0))$.*

A Jacobi field with these properties is called a $\Sigma$-*Jacobi field* along $c$.

*Proof*: Suppose first that (1) holds. Put $\sigma(s) := \gamma_s(0) \in \Sigma$ and $\Gamma(s) := \gamma_s'(0) \in T\Sigma^\perp_{\sigma(s)}$. Then $Y(0) = \sigma'(0) \in T\Sigma_p$. Moreover,

$$Y'(0) = \frac{D}{dt}\Big|_{t=0} Y = \frac{D}{ds}\Big|_{s=0} \Gamma.$$

Note that $\Gamma(0) = c'(0) = N(p) = (N \circ \sigma)(0)$. Hence, if $Z$ is any vector field along $\sigma$ tangent to $\Sigma$, then $\big\langle \frac{D}{ds}(\Gamma - N \circ \sigma), Z\big\rangle = -\big\langle \Gamma - N \circ \sigma, \frac{D}{ds} Z\big\rangle$ vanishes at $s = 0$. It follows that

$$\Big(\frac{D}{ds}\Big|_{s=0}\Gamma\Big)^{\mathrm{T}} = \Big(\frac{D}{ds}\Big|_{s=0}(N \circ \sigma)\Big)^{\mathrm{T}} = (D_{Y(0)}N)^{\mathrm{T}}.$$

Combining these equalities we get (2).

Conversely, suppose that (2) holds. Choose a curve $\sigma\colon (-\epsilon,\epsilon) \to \Sigma$ with $\sigma'(0) = Y(0) \in T\Sigma_p$. Let $X$ and $W$ be any vector fields along $\sigma$ normal to $\Sigma$ such that $X(0) = N(p)$ and $W(0)$ equals the normal component of $Y'(0) - \frac{D}{ds}\big|_{s=0}X$. Put $\Gamma(s) := X(s) + sW(s)$. Note that $\Gamma(0) = N(p) = c'(0)$. Similarly as above, $\big(\frac{D}{ds}\big|_{s=0}X\big)^{\mathrm{T}} = (D_{Y(0)}N)^{\mathrm{T}}$, thus

$$\frac{D}{ds}\Big|_{s=0}\Gamma = \frac{D}{ds}\Big|_{s=0}X + W(0) = (D_{Y(0)}N)^{\mathrm{T}} + Y'(0)^{\perp} = Y'(0)$$

by (2). Now it follows as in the proof of Proposition 3.6 that (1) holds. $\square$

**3.21 Definition** The point $q = c(l)$ is called a *focal point* of $\Sigma$ along $c$ if there exists a non-trivial $\Sigma$-Jacobi field $Y$ along $c$ with $Y(l) = 0$.

In analogy to Lemma 3.11 it can be shown that $q$ is focal point of $\Sigma$ along $c$ if and only if $lc'(0)$ is a singular point of the *normal exponential map* $\exp^{\perp} := \exp|_{\Omega\cap T\Sigma^{\perp}}$.

**3.22 Remark** Suppose now that $\dim(\Sigma) = \dim(M) - 1$ and $S_{c'(0)} = 0$. (This holds, for example, if $\Sigma$ is the *geodesic submanifold* $\exp_p(c'(0)^{\perp} \cap B(r))$ orthogonal to $c'(0)$, for some sufficiently small $r > 0$.) Then it follows from Proposition 3.20 that a normal Jacobi field $Y$ along $c$ is a $\Sigma$-Jacobi field if and only if $Y'(0) = 0$. Hence, $q = c(l)$ is a focal point of $\Sigma$ along $c$ if and only if there exists a non-trivial Jacobi field $Y$ along $c$ with $Y'(0) = 0$ and $Y(l) = 0$ (which is then necessarily normal). If *no* such Jacobi fields exists, then the linear map that sends $v \in TM_p$ to $Y_v(l) \in TM_q$ for the Jacobi field $Y_v$ with $Y_v(0) = v$ and $Y_v'(0) = 0$ is an isomorphism, thus for every $w \in TM_q$ there is a unique Jacobi field $Y$ along $c$ such that $Y'(0) = 0$ and $Y(l) = w$. If $w$ is normal, then so is $Y$.

**3.23 Proposition (second index lemma)** *Let $c\colon [0,l] \to M$ be a unit speed geodesic, and suppose that no non-trivial Jacobi field $Y$ along $c$ with $Y'(0) = 0$ has a zero in $(0,l]$. If $X$ is a piecewise smooth normal vector field along $c$ and $Y$ is the unique Jacobi field along $c$ such that $Y'(0) = 0$ and $Y(l) = X(l)$, then*

$$I(X,X) \geq I(Y,Y),$$

*and equality holds if and only if $X = Y$.*

*Proof*: Let $V_1,\dots,V_{m-1}$ be Jacobi fields along $c$ such that $(V_1(0),\dots,V_{m-1}(0))$ is an orthonormal basis of $c'(0)^{\perp}$, and $V_i'(0) = 0$ for $k = 1,\dots,m-1$. It follows from the assumption on $c$ that $(V_1(t),\dots,V_{m-1}(t))$ is a basis of $c'(t)^{\perp}$ for every $t \in [0,l]$. Thus there exist piecewise smooth functions $\lambda_1,\dots,\lambda_{m-1}\colon [0,l] \to \mathbb{R}$ such that $X = \sum_i \lambda_i V_i$. The linear combination $\sum_i \lambda_i(l)V_i$ (with constant coefficients) is a

Jacobi field with covariant derivative 0 at $t = 0$ and value $X(l) = Y(l)$ at $t = l$, so it is equal to $Y$. Since $Y'(0) = 0$,

$$I(Y,Y) = \langle Y(l), Y'(l)\rangle = \sum_{i=1}^{m-1} \bigl\langle X(l), \lambda_i(l) V_i'(l)\bigr\rangle.$$

As $V_i'(0) = 0$, the summand $\langle X, \lambda_i V_i'\rangle(l)$ can be rewritten as the integral over $[0,l]$ of the derivative of the piecewise smooth function $\langle X, \lambda_i V_i'\rangle$. Together with the Jacobi equation for $V_i$, this gives

$$\begin{aligned}\langle X, \lambda_i V_i'\rangle(l) &= \int_0^l \langle X', \lambda_i V_i'\rangle + \langle X, \lambda_i' V_i'\rangle + \langle X, \lambda_i V_i''\rangle \, dt\\ &= \int_0^l \langle X', (\lambda_i V_i)'\rangle - \lambda_i' f_i - \langle X, R(\lambda_i V_i, c')c'\rangle \, dt,\end{aligned}$$

where $f_i := \langle X', V_i\rangle - \langle X, V_i'\rangle$. Taking again the sum, we get that

$$I(Y,Y) = I(X,X) - \sum_{i=1}^{m-1} \int_0^l \lambda_i' f_i \, dt.$$

Now $f_i = \sum_j \lambda_j' \langle V_j, V_i\rangle + \lambda_j \bigl(\langle V_j', V_i\rangle - \langle V_j, V_i'\rangle\bigr)$, and the function $\langle V_j', V_i\rangle - \langle V_j, V_i'\rangle$ vanishes at $t = 0$ and has zero derivative on $[0,l]$, as one readily checks from the Jacobi equation and the symmetry properties of $R$. It follows that

$$I(Y,Y) = I(X,X) - \int_0^l \left| \sum_{i=1}^{m-1} \lambda_i' V_i \right|^2 dt \le I(X,X).$$

Equality holds if and only if every $\lambda_i$ is constant, that is, $X = \sum_i \lambda_i(l) V_i = Y$. □

**3.24 Theorem (Berger 1962)** *Let again $M, \bar M$ and $c, \bar c$ be given as in Assumptions* 3.17*, and suppose that no non-trivial Jacobi field $\bar Y$ along $\bar c$ with $\bar Y'(0) = 0$ has a zero in* $(0,l)$*. If $Y$ and $\bar Y$ are Jacobi fields along $c$ and $\bar c$, respectively, such that $Y'(0) = 0 = \bar Y'(0)$ and*

$$|Y(0)| = |\bar Y(0)|, \quad \langle Y(0), c'(0)\rangle = \langle \bar Y(0), \bar c\,'(0)\rangle,$$

*then $|Y(t)| \ge |\bar Y(t)|$ for all $t \in [0,l]$.*

*Proof*: Suppose, as in the proof of Theorem 3.18, that $Y, \bar Y$ are normal along $c, \bar c$ and $\bar Y$ is non-trivial. Then it follows from the assumptions that $\bar Y(t) \ne 0$ for all $t \in [0,l)$. Now the same argument as in the proof of Theorem 3.18, with Proposition 3.23 in place of Proposition 3.16, shows again that $\bigl(|Y|^2/|\bar Y|^2\bigr)' \ge 0$ on $(0,l)$. Since $|Y(0)| = |\bar Y(0)| \ne 0$, this yields the result. □

# Chapter 4

# Riemannian Submersions and Coverings

## Riemannian submersions

**4.1 Definition** Let $(\bar{M}, \bar{g})$, $(M, g)$ be two Riemannian manifolds, and suppose that $\pi \colon \bar{M} \to M$ is a submersion (that is, $d\pi_p \colon T\bar{M}_p \to TM_{\pi(p)}$ is surjective for all $p \in \bar{M}$). Then $\pi$ is called a *Riemannian submersion* if for all $p \in \bar{M}$ the map

$$d\pi_p|_{H_p} \colon H_p \to TM_{\pi(p)}$$

is an isometry, where $H_p \subset T\bar{M}_p$ is the orthogonal complement of $V_p := \ker(d\pi_p)$. Vectors in $H_p$ or $V_p$ are called *horizontal* or *vertical*, respectively.

Note that for every $p \in \pi(\bar{M}) \subset M$, the fiber $\pi^{-1}\{p\}$ is a submanifold of $\bar{M}$ of dimension $\dim(\bar{M}) - \dim(M)$.

**4.2 Lemma** *If $\bar{M}$ and $M$ are two connected Riemannian manifolds with distance functions $\bar{d}$ and $d$, and if $\pi \colon \bar{M} \to M$ is a Riemannian submersion, then $d(\pi(p), \pi(q)) \le \bar{d}(p, q)$ for all $p, q \in \bar{M}$.*

*Proof*: For a vector $v \in T\bar{M}_p$, let $v = v^{\mathrm{hor}} + v^{\mathrm{ver}}$ be its decomposition into horizontal and vertical part. Since $|d\pi_p(v)| = |d\pi_p(v^{\mathrm{hor}})| = |v^{\mathrm{hor}}| \le |v|$, it follows that $L(\gamma \circ \pi) \le L(\gamma)$ for every piecewise $C^1$ curve in $\bar{M}$. □

The following result relates geodesics in $M$ to horizontal geodesics in $\bar{M}$.

**4.3 Proposition** *Suppose that $\pi \colon \bar{M} \to M$ is a Riemannian submersion.*

(1) *Let $c$ be a maximal geodesic in $M$ with $c(0) = \pi(p)$ for some $p \in M$. Then there exists a unique maximal horizontal curve $\bar{c} \colon I \to \bar{M}$ with $\bar{c}(0) = p$ and $\pi \circ \bar{c} = c|_I$, and $\bar{c}$ is a geodesic in $\bar{M}$.*

(2) *Let $\bar{c}$ be a geodesic in $\bar{M}$ with horizontal initial velocity $\bar{c}\,'(0)$. Then $\bar{c}$ is horizontal, and $c := \pi \circ \bar{c}$ is a geodesic in $M$.*

Notice that if $\bar{M}$ is complete and $M$ is connected, then it follows from (2) and Theorem 1.23 (Hopf–Rinow) that $M$ is complete and $\pi$ is surjective.

*Proof*: We prove (1). Assume that $c'(0) \neq 0$. Let $\epsilon > 0$ be such that $N := c((-\epsilon, \epsilon))$ is a 1-dimensional submanifold of $M$ contained in $\pi(\bar{M})$. As a submersion, $\pi$ is transverse to $N$, and it follows from a standard result in differential topology that $\pi^{-1}(N)$ is a submanifold of $\bar{M}$ of dimension $\dim(\bar{M}) - \dim(M) + 1$. Consider the horizontal lift $\bar{X}$ of $X := c'|_N$ on $\pi^{-1}(N)$. Let $\bar{c}$ be the (maximal) integral curve of $\bar{X}$ with $\bar{c}(0) = p$. For every $t$ in its domain,

$$(\pi \circ \bar{c})'(t) = d\pi_{\bar{c}(t)}(\bar{c}\,'(t)) = d\pi_{\bar{c}(t)}(\bar{X}_{\bar{c}(t)}) = X_{\pi\circ\bar{c}(t)},$$

thus $\pi \circ \bar{c}$ is the integral curve of $X$ with $\pi \circ \bar{c}(0) = c(0)$ and so $\pi \circ \bar{c} = c|_{(-\epsilon,\epsilon)}$. Then $\bar{c}$ is a geodesic since $|\bar{c}\,'(t)| = |\bar{X}_{\bar{c}(t)}| = |X_{c(t)}|$ is independent of $t$ and, for suitable subintervals $[t, t']$,

$$L(\bar{c}|_{[t,t']}) = L(c|_{[t,t']}) = d(c(t), c(t')) \leq d(\bar{c}(t), \bar{c}(t'))$$

by Lemma 4.2. Now (1) follows.

For the proof of (2), given $\bar{c} \colon I \to \bar{M}$, let $\gamma$ be the maximal geodesic in $M$ with $\gamma'(0) = d\pi_{\bar{c}(0)}(\bar{c}\,'(0))$. By (1), there is a maximal horizontal lift $\bar{\gamma}$ with $\bar{\gamma}(0) = \bar{c}(0)$, and $\bar{\gamma}$ is a geodesic in $\bar{M}$. Since $\bar{c}\,'(0)$ is horizontal, $\bar{c}\,'(0) = \bar{\gamma}'(0)$, thus $\bar{c} = \bar{\gamma}|_I$ is horizontal, and $c := \pi \circ \bar{c} = \pi \circ \bar{\gamma}|_I = \gamma|_I$. □

**Example** Let $(M, g)$ and $(M', g')$ be Riemannian manifolds, and let $f \colon M \to (0, \infty)$ be a $C^\infty$ function. The *warped product* $M \times_f M'$ is the product manifold $M \times M'$ endowed with the metric defined by

$$(g \times_f g')_{(p,p')}((v, v'), (w, w')) = g_p(v, w) + f(p)^2 g'_{p'}(v', w')$$

for $(v, v'), (w, w') \in T(M \times M')_{(p,p')} = TM_p \times TM'_{p'}$; briefly, $g \times_f g' = g + f^2 g'$. Then the canonical projection

$$\pi \colon (M \times M', g \times_f g') \to (M, g)$$

is a Riemannian submersion with

$$V_{(p,p')} = \{0_p\} \times TM'_{p'} \quad \text{and} \quad H_{(p,p')} = TM_p \times \{0_{p'}\}.$$

For every $p' \in M'$, the submanifold $M \times \{p'\}$ is everywhere horizontal and therefore totally geodesic in $M \times_f M'$ (recall Proposition 2.18).

**4.4 Proposition** *Let $\pi\colon \bar{M} \to M$ be a surjective submersion, and let $\bar{g}$ be a Riemannian metric on $\bar{M}$. Suppose that for every pair of points $p, q \in \bar{M}$ with $\pi(p) = \pi(q)$ there exists an isometry $h \in \mathrm{Isom}(\bar{M}, \bar{g})$ such that $h(p) = q$ and $\pi \circ h = \pi$. Then there exists a unique Riemannian metric $g$ on $M$ such that $\pi\colon (\bar{M}, \bar{g}) \to (M, g)$ is a Riemannian submersion.*

*Proof*: First we verify that for every pair of points $p, q \in \bar{M}$ with $\pi(p) = \pi(q)$, the map

$$(d\pi_q|_{H_q})^{-1} \circ (d\pi_p|_{H_p})\colon H_p \to H_q \tag{1}$$

is an isometry. By assumption there exists an isometry $h$ of $\bar{M}$ such that $h(p) = q$ and $\pi \circ h = \pi$. Then $d\pi_q \circ dh_p = d\pi_p$, and thus the isomorphism $dh_p$ maps $V_p = \ker(d\pi_p)$ to $V_q$. Hence, as an isometry, $dh_p$ maps $H_p$ isometrically to $H_q$, and the map (1) is just $dh_p|_{H_p}$. Since $\pi$ is surjective, it now follows that for every $x \in M$ there is a unique inner product $g_x$ on $TM_x$ such that $d\pi_p|_{H_p}: H_p \to TM_x$ is an isometry for all $p \in \pi^{-1}\{x\}$. It can be shown that $g_x$ depends smoothly on $x$. □

**Example** The *complex projective space* $\mathbb{C}P^n$ may be regarded as the quotient space of the unit sphere $S^{2n+1} \subset \mathbb{R}^{2n+2} = \mathbb{C}^{n+1}$ by the equivalence relation

$$z \sim z' \iff z' = \lambda z \text{ for some } \lambda = \mathrm{e}^{i\theta} \in S^1 \subset \mathbb{C}.$$

Then the quotient projection $\pi\colon S^{2n+1} \to \mathbb{C}P^n$ is a surjective submersion, the *(generalized) Hopf fibration*. For any $\lambda$ as above, the map $h_\lambda\colon S^{2n+1} \to S^{2n+1}$ sending $z$ to $\lambda z$ (a rotation) is an isometry of $(S^{2n+1}, g^{\mathrm{sph}})$ satisfying $\pi \circ h_\lambda = \pi$. By Proposition 4.4 there exists a unique Riemannian metric $g$ on $\mathbb{C}P^n$ with respect to which $\pi$ is a Riemannian submersion, called the *Fubini–Study metric*.

For $p \in S^{2n+1}$ and a unit vector $v \in TS^{2n+1}_p$, let $\bar{c} = \bar{c}_v\colon \mathbb{R} \to S^{2n+1}$ be the geodesic given by $\bar{c}(t) = p\cos(t) + v\sin(t)$. Note that the fiber of $\pi$ through $p$ is contained in the plane with orthonormal basis $\{p, ip\}$, thus the vector $ip \in TS^{2n+1}_p$ is tangent to the fiber, and $\bar{c}$ is horizontal if and only if $\langle ip, v\rangle = 0$. If this condition holds, then $c := \pi \circ \bar{c}$ is the geodesic in $\mathbb{C}P^n$ with $c(0) = \pi(p)$ and $c'(0) = d\pi_p(v)$ (by Proposition 4.3). For all $t \in \mathbb{R}$,

$$c(t+\pi) = \pi(\bar{c}(t+\pi)) = \pi(-\bar{c}(t)) = \pi(\bar{c}(t)) = c(t),$$

thus $c$ is periodic with period $\pi$.

## Curvature of Riemannian submersions

Let $\pi\colon \bar{M} \to M$ be a Riemannian submersion. We now discuss some results from [ON1966] relating the Levi-Civita connections $\bar{D}$ and $D$, as well as the curvature tensors $\bar{R}$ and $R$, of $\bar{M}$ and $M$. For any vector field $X \in \Gamma(TM)$, we let $\overline{X} \in \Gamma(T\bar{M})$ denote the unique horizontal vector field such that $\pi_* \overline{X} = X \circ \pi$ (that is, $\overline{X}$ and $X$ are *$\pi$-related*). Notice that for all $X, Y \in \Gamma(TM)$,

$$\langle \overline{X}, \overline{Y} \rangle = \langle \pi_* \overline{X}, \pi_* \overline{Y} \rangle = \langle X, Y \rangle \circ \pi \quad \text{and} \quad \pi_*[\overline{X}, \overline{Y}] = [X, Y] \circ \pi.$$

**4.5 Proposition** *For all $X, Y, Z \in \Gamma(TM)$,*

$$\left\langle \bar{D}_{\overline{X}} \overline{Y}, \overline{Z} \right\rangle = \langle D_X Y, Z \rangle \circ \pi \quad \textit{and} \quad \bar{D}_{\overline{X}} \overline{Y} = \overline{D_X Y} + \frac{1}{2} [\overline{X}, \overline{Y}]^{\mathrm{ver}}.$$

*Proof*: For all $X, Y, Z \in \Gamma(TM)$,

$$\overline{Z} \langle \overline{X}, \overline{Y} \rangle = \overline{Z}(\langle X, Y \rangle \circ \pi) = (\pi_* \overline{Z}) \langle X, Y \rangle = (Z, \langle X, Y \rangle) \circ \pi$$

and

$$\langle \overline{Z}, [\overline{X}, \overline{Y}] \rangle = \langle \pi_* \overline{Z}, \pi_* [\overline{X}, \overline{Y}] \rangle = \langle Z, [X, Y] \rangle \circ \pi.$$

The assertion $\left\langle \bar{D}_{\overline{X}} \overline{Y}, \overline{Z} \right\rangle = \langle D_X Y, Z \rangle \circ \pi$ follows readily from these relations and the Koszul formulas for $\bar{D}$ and $D$. In turn, $\langle D_X Y, Z \rangle \circ \pi = \left\langle \overline{D_X Y}, \overline{Z} \right\rangle$, thus

$$\left( \bar{D}_{\overline{X}} \overline{Y} \right)^{\mathrm{hor}} = \overline{D_X Y}. \tag{1}$$

Now let $T \in \Gamma(T\bar{M})$ be a vertical vector field. Then $\pi_* T = 0$ and $\pi_*[\overline{X}, T] = [X, 0] \circ \pi = 0$, so $[\overline{X}, T]$ is vertical, and the Koszul formula for $\bar{D}$ yields

$$2 \left\langle \bar{D}_{\overline{X}} \overline{Y}, T \right\rangle = -T \langle \overline{X}, \overline{Y} \rangle + \langle T, [\overline{X}, \overline{Y}] \rangle,$$

where $T \langle \overline{X}, \overline{Y} \rangle = T(\langle X, Y \rangle \circ \pi) = (\pi_* T) \langle X, Y \rangle = 0$. Hence

$$\left( \bar{D}_{\overline{X}} \overline{Y} \right)^{\mathrm{ver}} = \frac{1}{2} [\overline{X}, \overline{Y}]^{\mathrm{ver}}, \tag{2}$$

and (1) and (2) yield the second assertion. □

**4.6 Theorem (O'Neill 1966)** *For all $X, Y \in \Gamma(TM)$,*

$$R(X, Y, X, Y) \circ \pi - \bar{R}(\overline{X}, \overline{Y}, \overline{X}, \overline{Y}) = \frac{3}{4} \left| [\overline{X}, \overline{Y}]^{\mathrm{ver}} \right|^2 = 3 \left| \left( \bar{D}_{\overline{X}} \overline{Y} \right)^{\mathrm{ver}} \right|^2.$$

*Proof*: Let $V, W, X, Y \in \Gamma(TM)$. From Proposition 4.5 we get that

$$\overline{X} \left\langle \overline{V}, \bar{D}_{\overline{Y}} \overline{W} \right\rangle = \overline{X}(\langle V, D_Y W \rangle \circ \pi) = (\pi_* \overline{X}) \langle V, D_Y W \rangle = (X \langle V, D_Y W \rangle) \circ \pi$$

and, for $A := \frac{1}{4} \langle [\overline{X}, \overline{V}]^{\mathrm{ver}}, [\overline{Y}, \overline{W}]^{\mathrm{ver}} \rangle$,

$$\left\langle \bar{D}_{\overline{X}} \overline{V}, \bar{D}_{\overline{Y}} \overline{W} \right\rangle = \left\langle \overline{D_X V}, \overline{D_Y W} \right\rangle + A = \langle D_X V, D_Y W \rangle \circ \pi + A.$$

It follows that

$$\text{(1)} \qquad \begin{aligned}\left\langle \overline{V}, \bar{D}_{\overline{X}} \bar{D}_{\overline{Y}} \overline{W} \right\rangle &= \overline{X}\left\langle \overline{V}, \bar{D}_{\overline{Y}} \overline{W} \right\rangle - \left\langle \bar{D}_{\overline{X}} \overline{V}, \bar{D}_{\overline{Y}} \overline{W} \right\rangle \\ &= (X\langle V, D_Y W\rangle) \circ \pi - \langle D_X V, D_Y W\rangle \circ \pi - A \\ &= \langle V, D_X D_Y W\rangle \circ \pi - A.\end{aligned}$$

Likewise, for $B := \frac{1}{4}\langle [\overline{Y}, \overline{V}]^{\mathrm{ver}}, [\overline{X}, \overline{W}]^{\mathrm{ver}}\rangle$,

$$\text{(2)} \qquad \left\langle \overline{V}, \bar{D}_{\overline{Y}} \bar{D}_{\overline{X}} \overline{W} \right\rangle = \langle V, D_Y D_X W\rangle \circ \pi - B.$$

Next, note that $\pi_*[\overline{X}, \overline{Y}]^{\mathrm{hor}} = [X, Y] \circ \pi$ and thus $[\overline{X}, \overline{Y}]^{\mathrm{hor}} = \overline{[X, Y]}$ by uniqueness. Write $[\overline{X}, \overline{Y}] = \overline{[X, Y]} + T$ for $T := [\overline{X}, \overline{Y}]^{\mathrm{ver}}$. Since $T$ is vertical, so is $[\overline{W}, T]$. From Proposition 4.5 we get that

$$\left\langle \overline{V}, \bar{D}_T \overline{W} \right\rangle = \left\langle \overline{V}, \bar{D}_{\overline{W}} T \right\rangle = -\left\langle \bar{D}_{\overline{W}} \overline{V}, T \right\rangle = \frac{1}{2}\langle [\overline{V}, \overline{W}]^{\mathrm{ver}}, T\rangle =: C$$

and $\left\langle \overline{V}, \bar{D}_{\overline{[X,Y]}} \overline{W} \right\rangle = \left\langle \overline{V}, \overline{D_{[X,Y]} W} \right\rangle$, hence

$$\text{(3)} \qquad \left\langle \overline{V}, \bar{D}_{[\overline{X}, \overline{Y}]} \overline{W} \right\rangle = \left\langle \overline{V}, \overline{D_{[X,Y]} W} \right\rangle + C = \langle V, D_{[X,Y]} W\rangle \circ \pi + C.$$

Now (1), (2), and (3) yield the identity

$$\bar{R}(\overline{V}, \overline{W}, \overline{X}, \overline{Y}) = R(V, W, X, Y) \circ \pi - A + B - C.$$

If $(V, W) = (X, Y)$, then $A = 0$ and $C - B = \frac{3}{4}\left|[\overline{X}, \overline{Y}]^{\mathrm{ver}}\right|^2 = 3\left|(\bar{D}_{\overline{X}} \overline{Y})^{\mathrm{ver}}\right|^2$. □

**Example** Consider again the Hopf fibration $\pi\colon S^{2n+1} \to \mathbb{C}P^n$, where $\mathbb{C}P^n$ is equipped with the Fubini–Study metric. Let $T \in \Gamma(TS^{2n+1})$ be the vertical unit vector field given by $T(p) = ip$. For $X, Y \in \Gamma(T(\mathbb{C}P^n))$ and the horizontal lifts $\overline{X}, \overline{Y} \in \Gamma(TS^{2n+1})$,

$$\left|(\bar{D}_{\overline{Y}} \overline{X})^{\mathrm{ver}}\right| = \left|\left\langle \bar{D}_{\overline{Y}} \overline{X}, T \right\rangle\right| = \left|\left\langle \overline{X}, \bar{D}_{\overline{Y}} T \right\rangle\right|.$$

Here, the Levi-Civita connection $\bar{D}$ of $S^{2n+1}$ is just the tangential part of the usual directional derivative in $\mathbb{R}^{2n+2}$, and it follows from the definition of $T$ that $\left\langle \overline{X}, \bar{D}_{\overline{Y}} T \right\rangle = \langle \overline{X}, i\overline{Y}\rangle$. Hence, if $x, y \in T(\mathbb{C}P^n)_{\pi(p)}$ are orthonormal and $\bar{x}, \bar{y} \in TS^{2n+1}_p$ are the horizontal lifts, then Theorem 4.6 gives the explicit formula

$$\sec(\operatorname{span}\{x, y\}) = 1 + 3\langle \bar{x}, i\bar{y}\rangle^2.$$

In particular, the sectional curvature of $\mathbb{C}P^n$ takes values in $[1, 4]$.

## Riemannian coverings and space forms

We now turn to the equi-dimensional case. Let $\bar{M}$ and $M$ be two manifolds of the same dimension. A smooth map $F\colon \bar{M} \to M$ is a *local diffeomorphism* if every point in $\bar{M}$ has an open neighborhood that is mapped diffeomorphically onto an open set in $M$. By the inverse function theorem, this holds if and only if $dF_p : T\bar{M}_p \to TM_{F(p)}$ is bijective for every $p \in \bar{M}$. Given Riemannian metrics $\bar{g}$ and $g$ on $\bar{M}$ and $M$, respectively, a smooth map $F\colon \bar{M} \to M$ is a *local isometry* if $F^*g = \bar{g}$. Then $F$ is in particular a local diffeomorphism (and a Riemannian submersion with discrete fibers). Note that, as in Definition 4.1, we do not assume the map $F$ to be surjective. Evidently, a local isometry takes geodesics to geodesics (compare Proposition 4.3).

**4.7 Lemma** *Suppose that $F, G\colon \bar{M} \to M$ are two local isometries, and $\bar{M}$ is connected. If $F(p) = G(p)$ and $dF_p = dG_p$ for some point $p \in \bar{M}$, then $F = G$.*

*Proof*: The set

$$A := \{q \in \bar{M} : F(q) = G(q) \text{ and } dF_q = dG_q\}$$

is non-empty (as $p \in A$) and closed because $F$ and $G$ are continuously differentiable. Let $q \in A$. Since $F$ and $G$ are local isometries, in a neighborhood of $0 \in T\bar{M}_q$,

$$F \circ \exp_q = \exp_{F(q)} \circ dF_q = \exp_{G(q)} \circ dG_q = G \circ \exp_q .$$

Hence $F = G$ in a neighborhood of $q$. This shows that $A$ is also open. Since $\bar{M}$ is connected, it follows that $A = \bar{M}$, in particular $F = G$ on $\bar{M}$. □

Next, recall that a *covering map* $\pi\colon \bar{M} \to M$ of topological spaces $\bar{M}, M$ is a continuous surjective map such that every point in $M$ has an open neighborhood $U$ whose preimage $\pi^{-1}(U)$ is a union of pairwise disjoint open sets each of which is mapped homeomorphically onto $U$ by $\pi$.

**4.8 Definition** Let again $(\bar{M}, \bar{g})$ and $(M, g)$ be two Riemannian manifolds of the same dimension. A smooth covering map $\pi\colon (\bar{M}, \bar{g}) \to (M, g)$ with the property that $\pi^*g = \bar{g}$ is called a *Riemannian covering map*.

**4.9 Proposition** *Suppose that $\bar{M}$ is a complete Riemannian manifold and $M$ is a connected Riemannian manifold of the same dimension. Then every local isometry $F\colon \bar{M} \to M$ is a Riemannian covering map.*

*Proof*: It follows from the assumptions that $M$ is complete and $F$ is surjective (compare the remark after Proposition 4.3). Now let $q \in M$, and choose $r > 0$ such that $\exp_q\colon B_r \to B(q, r)$ is a diffeomorphism.

First we show that for every pair of distinct points $p, p' \in F^{-1}\{q\}$, the (open) balls $B(p,r)$ and $B(p',r)$ are disjoint. Since $\bar{M}$ is complete there exists a minimizing geodesic $\bar{c}$ from $p$ to $p'$. Furthermore, since $F$ is a local isometry, $c := F \circ \bar{c}$ is a geodesic loop at $q$. Then $d(p,p') = L(\bar{c}) = L(c) \geq 2r$ by the choice of $r$.

Next we show that $F|_{B(p,r)} \colon B(p,r) \to B(q,r)$ is a diffeomorphism for every $p \in F^{-1}\{q\}$. Since $F$ maps geodesics to geodesics,

$$F \circ \exp_p |_{B_r} = \exp_q \circ \, dF_p|_{B_r},$$

and the latter is a diffeomorphism onto $B(q,r)$ because $dF_p \colon TM_p \to TM_q$ is an isometry and due to the choice of $r$. This yields the result.

Lastly, $\bigcup_{p \in F^{-1}\{q\}} B(p,r) \subset F^{-1}(B(q,r))$, because $F(B(p,r)) \subset B(q,r)$ for all $p \in F^{-1}\{q\}$. For the reverse inclusion, let $x \in F^{-1}(B(q,r))$. Since $d(F(x),q) < r$, there is a vector $w \in TM_{F(x)}$ such that $|w| < r$ and $\exp_{F(x)}(w) = q$. Let $v := (dF_x)^{-1}(w) \in T\bar{M}_x$ be its lift. Then

$$F(\exp_x(v)) = \exp_{F(x)}(dF_x(v)) = \exp_{F(x)}(w) = q,$$

thus $p := \exp_x(v) \in F^{-1}\{q\}$. As $|v| = |w| < r$, it follows that $x \in B(p,r)$. $\square$

Let, for the moment, $\bar{M}$ be a connected *topological* manifold. A group $\Gamma \subset \mathrm{Homeo}(\bar{M})$ of homeomorphisms of $\bar{M}$ acts *freely* on $\bar{M}$ if $\gamma(p) \neq p$ whenever $\gamma \in \Gamma \setminus \{\mathrm{id}\}$ and $p \in \bar{M}$, and $\Gamma$ acts *properly discontinuously* if for every compact set $K \subset \bar{M}$ there are only finitely many elements $\gamma \in \Gamma$ with $\gamma(K) \cap K \neq \emptyset$. If both properties hold, then the quotient space $\bar{M}/\Gamma$ is a topological manifold and the projection $\pi \colon \bar{M} \to \bar{M}/\Gamma$ is a covering map. Conversely, suppose that $F \colon \bar{M} \to M$ is a covering map onto another topological manifold $M$ (it suffices to assume that $M$ is a Hausdorff space). Then a homeomorphism $\gamma \in \mathrm{Homeo}(\bar{M})$ is called a *deck transformation* or *covering transformation* if $F \circ \gamma = F$, and the group $\Gamma$ of all deck transformations acts freely and properly discontinuously on $\bar{M}$. (See, for example, the proof of Proposition 3.5.7 in [Th1997].) The bijection sending each orbit $\Gamma(p)$ to $F(p)$ is a canonical homeomorphism from $\bar{M}/\Gamma$ onto $M$. Recall also that if $\bar{M}$ is simply connected, then $\Gamma$ is isomorphic to the fundamental group $\pi_1(M)$.

Let now again $\bar{M}$ be a connected *smooth* manifold, and let $\Gamma \subset \mathrm{Diff}(\bar{M})$ be a group of ($C^\infty$) diffeomorphisms of $\bar{M}$ that acts freely and properly discontinuously. Then there is a unique $C^\infty$ structure on $\bar{M}/\Gamma$ such that the covering map $\pi \colon \bar{M} \to \bar{M}/\Gamma$ is a local diffeomorphism. Furthermore, if $\Gamma \subset \mathrm{Isom}(\bar{M})$ is a group of isometries with respect to a Riemannian metric $\bar{g}$ on $\bar{M}$, then there is a unique Riemannian metric $g$ on $\bar{M}/\Gamma$ such that $\pi$ is a local isometry (compare Proposition 4.4).

The following characterization of space forms (recall Definition 2.6) is due to Killing [Ki1891] and Hopf [Ho1926]. For $\kappa \in \mathbb{R}$ and $m \geq 2$ we let $\mathbb{M}^m_\kappa$ denote the

$m$-dimensional model space of constant sectional curvature $\kappa$,

$$\mathbb{M}^m_\kappa = \begin{cases} (S^m, \frac{1}{\kappa} g^{\mathrm{sph}}) & \text{if } \kappa > 0, \\ (\mathbb{R}^m, g^{\mathrm{eucl}}) & \text{if } \kappa = 0, \\ (H^m, \frac{1}{|\kappa|} g^{\mathrm{hyp}}) & \text{if } \kappa < 0. \end{cases}$$

We write $D_\kappa$ for the diameter of $\mathbb{M}^m_\kappa$, thus $D_\kappa = \frac{\pi}{\sqrt{\kappa}}$ if $\kappa > 0$ and $D_\kappa = \infty$ if $\kappa \leq 0$.

**4.10 Theorem (Killing 1891, Hopf 1926)** *Let $M$ be an $m$-dimensional space form of curvature $\kappa \in \mathbb{R}$. Then there exists a group $\Gamma \subset \mathrm{Isom}(\mathbb{M}^m_\kappa)$ that acts freely and properly discontinuously on $\mathbb{M}^m_\kappa$ such that $M$ is isometric to $\mathbb{M}^m_\kappa/\Gamma$. If $M$ is simply connected, then $M$ isometric to $\mathbb{M}^m_\kappa$.*

*Proof*: Choose points $p \in \mathbb{M}^m_\kappa$ and $q \in M$, a linear isometry $H\colon T(\mathbb{M}^m_\kappa)_p \to TM_q$, and define

$$F := \exp_q \circ H \circ \left(\exp_p |_{B(D_\kappa)}\right)^{-1} \colon B(p, D_\kappa) \to M.$$

Since both $\mathbb{M}^m_\kappa$ and $M$ have constant curvature $\kappa$, it follows from Corollary 3.19 that $F$ is a local isometry.

Consider first the case $\kappa \leq 0$. Then $F$ is defined on all of $\mathbb{M}^m_\kappa$, and $F\colon \mathbb{M}^m_\kappa \to M$ is a covering map according to Proposition 4.9. The group $\Gamma \subset \mathrm{Homeo}(\mathbb{M}^m_\kappa)$ of deck transformations of $F$ acts freely and properly discontinuously on $\mathbb{M}^m_\kappa$. In fact, $\Gamma \subset \mathrm{Diff}(\mathbb{M}^m_\kappa)$, because $F$ is a local diffeomorphism, and there is an induced $C^\infty$ structure on $\mathbb{M}^m_\kappa/\Gamma$ such that the projection $\pi\colon \mathbb{M}^m_\kappa \to \mathbb{M}^m_\kappa/\Gamma$ is a local diffeomorphism and $\mathbb{M}^m_\kappa/\Gamma$ is diffeomorphic to $M$. For $\gamma \in \Gamma$ and $p' \in \mathbb{M}^m_\kappa$, differentiation of $F \circ \gamma = F$ gives $dF_{\gamma(p')} \circ d\gamma_{p'} = dF_{p'}$. Since $F$ is a local isometry, it follows that $\Gamma \subset \mathrm{Isom}(\mathbb{M}^m_\kappa)$, and $\mathbb{M}^m_\kappa/\Gamma$ carries a Riemannian metric such that $\pi$ is a local isometry. Thus $M$ is isometric to $\mathbb{M}^m_\kappa/\Gamma$.

Secondly, suppose that $\kappa > 0$. Choose a point $\tilde{p} \in \mathbb{M}^m_\kappa \setminus \{p, -p\}$ and define $\tilde{q} := F(\tilde{p})$, $\tilde{H} := dF_{\tilde{p}}$, and

$$\tilde{F} := \exp_{\tilde{q}} \circ \tilde{H} \circ \left(\exp_{\tilde{p}} |_{B(D_\kappa)}\right)^{-1} \colon B(\tilde{p}, D_\kappa) \to M.$$

Then $\tilde{F}(\tilde{p}) = \tilde{q} = F(\tilde{p})$, $d\tilde{F}_{\tilde{p}} = \tilde{H} = dF_{\tilde{p}}$, and $\tilde{F}$ is a local isometry, like $F$. Now Lemma 4.7 shows that $F$ and $\tilde{F}$ agree on the intersection of their domains, $\mathbb{M}^m_\kappa \setminus \{-p, -\tilde{p}\}$, and therefore $F$ extends to a local isometry $F\colon \mathbb{M}^m_\kappa \to M$ with $F(-p) = \tilde{F}(-p)$. The rest of the argument is the same as in the case $\kappa \leq 0$. $\square$

**Examples**

1. (Euclidean space forms) Let $\Gamma \subset \mathrm{Isom}(\mathbb{R}^m, g^{\mathrm{eucl}})$ be a group of translations that acts freely and properly discontinuously. Then there exist linearly independent vectors $v_1, \ldots, v_k \in \mathbb{R}^m$ such that $\Gamma$ is the group

$$\Gamma = \left\{x \mapsto x + \textstyle\sum_{i=1}^k z_i v_i : (z_1, \ldots, z_k) \in \mathbb{Z}^k\right\}$$

isomorphic to $\mathbb{Z}^k$ (exercise). If $k = m$, then $(\mathbb{R}^m, g^{\mathrm{eucl}})/\Gamma$ is a *flat m-torus*, diffeomorphic to $T^m = \mathbb{R}^m/\mathbb{Z}^m$. If $k < m$, then $(\mathbb{R}^m, g^{\mathrm{eucl}})/\Gamma$ is isometric to the product $(T^k \times \mathbb{R}^{m-k}, g \times g^{\mathrm{eucl}})$ for some flat $k$-torus $(T^k, g)$.

If $\Gamma \subset \mathrm{Isom}(\mathbb{R}^m, g^{\mathrm{eucl}})$ acts freely and properly discontinuously with compact quotient, then the group $\Gamma' \subset \Gamma$ of translations has finite index in $\Gamma$ and $\mathbb{R}^m/\Gamma$ is finitely covered by the flat flat torus $\mathbb{R}^m/\Gamma'$. This is due to Bieberbach [Bi1911], [Bi1912]. See [Bus1985] for an elegant geometric proof.

2. (Hyperbolic space forms) Every compact oriented surface of genus $n \geq 2$ can be realized, in a flexible way, as a quotient $(H^2, g^{\mathrm{hyp}})/\Gamma$. The construction depends on $6n-6$ parameters (*Fenchel–Nielsen coordinates*), the corresponding moduli space is called *Teichmüller space*. (Oswald Teichmüller introduced quasiconformal mappings to the subject [Te1940].) See [FeN2003] for an edition of the Fenchel–Nielsen manuscript, and [Hu2006] for a detailed introduction to Teichmüller theory.

   For quotients of $(H^m, g^{\mathrm{hyp}})$ with $m \geq 3$ the famous *Mostow rigidity theorem* holds [Mo1968], [Mo1973]: if $H^m/\Gamma$ and $H^m/\Gamma'$ are compact and $\Gamma, \Gamma'$ are isomorphic, then $H^m/\Gamma$ and $H^m/\Gamma'$ are isometric.

3. (Spherical space forms) For the standard sphere $(S^m, g^{\mathrm{sph}})$ and $\Gamma = \{\mathrm{id}, -\mathrm{id}\}$,

$$(S^m, g^{\mathrm{sph}})/\Gamma = (\mathbb{R}P^m, g^{\mathrm{ell}})$$

   is the real projective space. The canonical (quotient) metric $g^{\mathrm{ell}}$ on $\mathbb{R}P^m$ is called the *elliptic metric*. For $m$ even, $S^m$ and $\mathbb{R}P^m$ are the only spherical space forms; see Theorem 4.11 below.

   For $m = 2n - 1$ odd, the sphere $S^m \subset \mathbb{R}^{m+1}$ admits other quotients. Choose integers $p, q_1, \ldots, q_n \geq 1$ such that $p$ and $q_j$ are coprime for $j = 1, \ldots, n$, and view $S^m$ as a subset of $\mathbb{C}^n$. Then the group of isometries

$$\Gamma := \left\{(z_1, \ldots, z_z) \mapsto (e^{2\pi i k q_1/p} z_1, \ldots, e^{2\pi i k q_n/p} z_n) : k = 0, \ldots, p-1\right\}$$

   isomorphic to $\mathbb{Z}/p\mathbb{Z}$ acts freely (and properly discontinuously, as a finite group) on $S^m$. The quotient $(S^m, g^{\mathrm{sph}})/\Gamma$ is the *lens space* $L(p; q_1, \ldots, q_n)$. Lens spaces were first described by Heinrich Tietze [Ti1908].

**4.11 Theorem** *Let $M$ be a space form with sectional curvature $\equiv 1$ and even dimension $m$. Then $M$ is isometric to $S^m$ or to the real projective space $\mathbb{R}P^m = S^m/\{\mathrm{id}, -\mathrm{id}\}$.*

*Proof*: From Theorem 4.10, we know that $M = S^m/\Gamma$ for some group $\Gamma \subset \mathrm{Isom}(S^m) \simeq O(m+1)$ that acts freely and properly discontinuously on $S^m$. Let

$\gamma \in \Gamma$. Since $m+1$ is odd, $\gamma$ has an eigenvalue 1 or $-1$. If $\gamma$ has an eigenvalue 1, then $\gamma = \mathrm{id}$ because $\Gamma$ acts freely. If $-1$ is the only real eigenvalue of $\gamma$, then 1 is an eigenvalue of $\gamma^2$ and hence $\gamma^2 = \mathrm{id}$. But then $\gamma = -\mathrm{id}$, for otherwise there would exist a vector $v \in S^m$ with $\gamma(v) \neq -v$, thus $\gamma(v) + v \neq 0$ would be an eigenvector with eigenvalue 1. This shows that $\Gamma = \{\mathrm{id}\}$ or $\Gamma = \{\mathrm{id}, -\mathrm{id}\}$. $\square$

A general reference for space forms is [Wo2011].

## Hadamard manifolds

The following result was established by Jacques Hadamard for surfaces and by Élie Cartan [Ca1928] in the general case.

**4.12 Theorem (Hadamard 1898, Cartan 1928)** *Let $M$ be a complete Riemannian manifold with sectional curvature* $\sec \leq 0$*, and let $p \in M$. Then* $\exp_p \colon TM_p \to M$ *is a covering map; in particular, if $M$ is simply connected, then* $\exp_p$ *is a diffeomorphism.*

A complete simply connected Riemannian manifold with non-positive sectional curvature is called a *Hadamard manifold* or a *Cartan–Hadamard manifold*.

*Proof*: Since $\sec \leq 0$, geodesics in $(M, g)$ have no conjugate points (exercise). Hence, by Lemma 3.11, $\exp_p \colon TM_p \to M$ has no singular points. It follows that $\bar{g} := \exp_p^* g$ defines a Riemannian metric on $TM_p$. For every $v \in TM_p$, the line $t \mapsto tv$ is a geodesic with respect to $\bar{g}$, because $t \mapsto \exp_p(tv)$ is a geodesic in $M$ and $\exp_p$ is a local isometry. Thus, by Theorem 1.23 (Hopf–Rinow), $(TM_p, \bar{g})$ is complete. Now Proposition 4.9 shows that $\exp_p \colon TM_p \to M$ is a covering map. $\square$

**4.13 Lemma** *Let $M$ be a Hadamard manifold. Then for every pair of geodesics $c, \bar{c} \colon \mathbb{R} \to M$, the function $t \mapsto d(c(t), \bar{c}(t))$ is convex on $\mathbb{R}$.*

*Proof*: Exercise. $\square$

**4.14 Proposition (flat strip)** *Suppose that $M$ is a Hadamard manifold, $c, \bar{c} \colon \mathbb{R} \to M$ are two unit speed geodesics with distinct images, and* $\sup_{s \in \mathbb{R}} d(c(s), \bar{c}(s)) < \infty$*. Let $f \colon \mathbb{R} \times [0,1] \to M$ be the geodesic homotopy between $c$ and $\bar{c}$; that is, for fixed $s \in \mathbb{R}$, $t \mapsto f(s,t)$ is the unique geodesic from $c(s)$ to $\bar{c}(s)$. Then there is an inner product on $\mathbb{R}^2$ with respect to which $f$ is a totally geodesic isometric embedding.*

The following argument (from [DeL2016]) is based entirely on Lemma 4.13.

*Proof*: For every $r \in \mathbb{R}$, the function $s \mapsto d(c(s), \bar{c}(s+r))$ is bounded and convex on $\mathbb{R}$ and thus equal to some constant $v(r)$. Furthermore $v(r) > 0$, for otherwise $c$

and $\bar{c}$ would have the same trace, contrary to the assumption. We now show that for every pair of points $x = (s, t)$ and $\bar{x} = (\bar{s}, t + \delta)$ in $\mathbb{R} \times [0, 1]$ with $\delta \geq 0$,

$$(1) \qquad d(f(x), f(\bar{x})) = \begin{cases} \delta\, \nu\big(\frac{\bar{s}-s}{\delta}\big) & \text{if } \delta > 0, \\ |\bar{s} - s| & \text{if } \delta = 0. \end{cases}$$

Suppose first that $\delta > 0$, and put $r := \frac{\bar{s}-s}{\delta}$. Let $y := (s - tr, 0)$ and $\bar{y} := (s + (1-t)r, 1)$ denote the points where the line through $x$ and $\bar{x}$ intersects $\mathbb{R} \times \{0\}$ and $\mathbb{R} \times \{1\}$. Then

$$(2) \qquad d(f(y), f(\bar{y})) = d(c(s - tr), \bar{c}(s - tr + r)) = \nu(r).$$

Write $\eta := f(s, \cdot\,)$ and $\bar{\eta} := f(\bar{s}, \cdot\,)$. By convexity,

$$(3) \qquad d(f(y), f(x)) = d(c(s - tr), \eta(t)) \leq t\, d(c(s - r), \eta(1)) = t\, \nu(r).$$

Similarly, we infer that

$$(4) \qquad d(f(\bar{x}), f(\bar{y})) \leq (1 - t - \delta)\, \nu(r)$$

as well as $d(\eta(0), \bar{\eta}(\delta)) \leq \delta\, \nu(r)$ and $d(\eta(1 - \delta), \bar{\eta}(1)) \leq \delta\, \nu(r)$. Hence, by the convexity of $\lambda \mapsto d(\eta(\lambda), \bar{\eta}(\lambda + \delta))$ on $[0, 1 - \delta]$, also

$$(5) \qquad d(f(x), f(\bar{x})) = d(\eta(t), \bar{\eta}(t + \delta)) \leq \delta\, \nu(r).$$

From (2)–(5) and the triangle inequality it follows that all inequalities derived so far are in fact equalities. In view of (5), this yields the first part of (1). For the second, we assume that the points $x = (s, t)$ and $\bar{x} = (\bar{s}, t)$ lie in $\mathbb{R} \times (0, 1)$. Putting $\bar{x}_\delta := (\bar{s}, t + \delta)$ for $\delta > 0$, we deduce that

$$d(f(x), f(\bar{x})) = \lim_{\delta \to 0+} d(f(x), f(\bar{x}_\delta)) = \lim_{\delta \to 0+} \delta\, \nu\Big(\frac{\bar{s} - s}{\delta}\Big).$$

Since $|r| - \nu(0) \leq \nu(r) \leq |r| + \nu(0)$ for all $r \in \mathbb{R}$, this limit is equal to $|\bar{s} - s|$, as required. It follows readily from (1) that there is a norm $\|\cdot\|$ on $\mathbb{R}^2$ such that $d(f(x), f(\bar{x})) = \|\bar{x} - x\|$ for all $x, \bar{x} \in \mathbb{R} \times [0, 1]$. Note that the triangle inequality for $\|\cdot\|$ is just inherited from $M$.

Finally, since $f$ is distance preserving with respect to $\|\cdot\|$, it follows that $\|v\|^2 = (f^*g)_x(v, v)$ for all $x \in \mathbb{R} \times [0, 1]$ and $v \in \mathbb{R}^2$. Thus the inner product $(f^*g)_x$ is independent of $x$ and induces $\|\cdot\|$. □

## Isometries of Hadamard manifolds

**4.15 Definition** Let $M$ be a Hadamard manifold, and let $\gamma \in \mathrm{Isom}(M)$ be an isometry. The *displacement function* $d_\gamma \colon M \to [0, \infty)$ is defined by $d_\gamma(p) := d(p, \gamma(p))$ for all $p \in M$. We put

$$|\gamma| := \inf\{d_\gamma(p) : p \in M\} \quad \text{and} \quad \mathrm{Min}(\gamma) := \{p \in M : d_\gamma(p) = |\gamma|\}.$$

**4.16 Lemma** *The set* $\mathrm{Min}(\gamma)$ *is closed, $\gamma$-invariant, and convex.*

*Proof*: It is clear that $\mathrm{Min}(\gamma)$ is closed. For every $p \in \mathrm{Min}(\gamma)$,

$$d_\gamma(\gamma p) = d(\gamma p, \gamma^2 p) = d(p, \gamma p) = d_\gamma(p) = |\gamma|,$$

thus $\mathrm{Min}(\gamma)$ is $\gamma$-invariant. Lastly, if $c\colon [0,1] \to M$ is a geodesic such that $c(0), c(1) \in \mathrm{Min}(\gamma)$, then it follows from Lemma 4.13 that

$$d(c(\lambda), \gamma c(\lambda)) \le (1 - \lambda)\, d(c(0), \gamma c(0)) + \lambda\, d(c(1), \gamma c(1)) = |\gamma|$$

for all $\lambda \in [0,1]$. This shows that $\mathrm{Min}(\gamma)$ is convex. □

Isometries of Hadamard manifolds are classified as follows.

**4.17 Definition** An isometry $\gamma$ of $M$ is called *parabolic* if $\mathrm{Min}(\gamma)$ is empty and *semi-simple* otherwise. In the latter case, $\gamma$ is *elliptic* if $|\gamma| = 0$ (that is, $\gamma$ has a fixed point) and *hyperbolic* if $|\gamma| > 0$.

For an isometry $\gamma$, a unit speed geodesic $c\colon \mathbb{R} \to M$ is called an *axis* of $\gamma$ if there exists a number $a > 0$ such that

$$\gamma(c(s)) = c(s + a) \quad \text{for all } s \in \mathbb{R}.$$

If $\gamma$ possesses an axis, then $\gamma$ will be called *axial*.

**4.18 Lemma** *Let* $\gamma \in \mathrm{Isom}(M)$*. If $\gamma$ has an axis $c$, then the corresponding number $a > 0$ is equal to $|\gamma|$, thus $\gamma$ is hyperbolic and* $c(\mathbb{R}) \subset \mathrm{Min}(\gamma)$*. Conversely, if $\gamma$ is hyperbolic, then for every point* $p \in \mathrm{Min}(\gamma)$ *there is an axis of $\gamma$ through $p$, and for every pair of axes $c, \bar{c}\colon \mathbb{R} \to M$ the function $s \mapsto d(c(s), \bar{c}(s))$ is constant.*

*Proof*: Suppose that $c$ is an axis of $\gamma$ with shift $a > 0$. For $p := c(0)$ and any $q \in M$, the triangle inequality gives

$$\begin{aligned} d(p, \gamma^n p) &\le d(p, q) + n\, d_\gamma(q) + d(\gamma^n q, \gamma^n p) \\ &= 2\, d(p, q) + n\, d_\gamma(q) \end{aligned}$$

for all $n \ge 1$, where $d(p, \gamma^n p) = na$. Thus $a \le d_\gamma(q)$ for all $q \in M$ and so $d_\gamma(p) = a = |\gamma|$. Hence $\gamma$ is hyperbolic, and $c(\mathbb{R}) \subset \mathrm{Min}(\gamma)$.

For the converse, let $\gamma$ be a hyperbolic isometry of $M$, put $a := |\gamma| > 0$, and let $p \in \mathrm{Min}(\gamma)$. Let $c\colon [0, a] \to M$ be the (unit speed) geodesic from $p$ to $\gamma p$, and extend it to a curve $c\colon \mathbb{R} \to M$ such that $c(na + t) = \gamma^n c(t)$ for all $n \in \mathbb{Z}$ and $t \in [0, a]$. Then, for all such $na + t =: s \in \mathbb{R}$,

$$\gamma\, c(s) = \gamma^{n+1} c(t) = c((n + 1)a + t) = c(s + a).$$

Since $c$ is parametrized by arc length, it follows that

$$a = L(c|_{[s,s+a]}) \geq d(c(s), c(s+a)) = d(c(s), \gamma\, c(s)) \geq |\gamma| = a.$$

Thus $c$ is a geodesic and an axis. If $\bar{c}\colon \mathbb{R} \to M$ is another axis of $\gamma$, then the function $s \mapsto d(c(s), \bar{c}(s))$ is bounded and convex, hence constant. $\square$

**4.19 Proposition** *Suppose that $M$ is Hadamard manifold, and $\Gamma$ is a subgroup of* $\mathrm{Isom}(M)$ *whose action is properly discontinuous and cocompact, that is, there is a compact set $K \subset M$ such that $\bigcup_{\alpha \in \Gamma} \alpha(K) = M$. Then every $\gamma \in \Gamma$ is semisimple* (*that is, hyperbolic or elliptic*). *In particular, if the action of $\Gamma$ is also free, then every $\gamma \in \Gamma \setminus \{\mathrm{id}\}$ is hyperbolic and hence axial.*

*Proof*: Let $\gamma \in \Gamma$. We want to show that the infimum $|\gamma| = \inf d_\gamma$ is attained. Let $p_1, p_2, \dots$ be a sequence of points in $M$ such that $d_\gamma(p_k) \to |\gamma|$ for $k \to \infty$. There exist a compact set $K \subset M$ and elements $\alpha_k \in \Gamma$ such that $q_k := \alpha_k^{-1} p_k \in K$ for all $k$. Put $\gamma_k := \alpha_k^{-1} \gamma \alpha_k$. Then

$$d_{\gamma_k}(q_k) = d(\alpha_k^{-1} p_k, \alpha_k^{-1} \gamma p_k) = d_\gamma(p_k),$$

hence the sets $\gamma_k(K)$ stay within bounded distance from $K$. Since $\Gamma$ acts properly discontinuously, it follows that the set $\{\gamma_k : k = 1, 2, \dots\}$ is finite. We now choose a sequence $k(i)$ such that all $\gamma_{k(i)}$ are equal to a fixed $\bar{\gamma} \in \Gamma$, and such that $q_{k(i)}$ converges to a point $q \in K$, as $i \to \infty$. Then $d_\gamma(\alpha_{k(i)} q) = d(q, \gamma_{k(i)} q) = d_{\bar{\gamma}}(q)$ for all $i$, and

$$d_{\bar{\gamma}}(q) = \lim_{i \to \infty} d_{\bar{\gamma}}(q_{k(i)}) = \lim_{i \to \infty} d_\gamma(p_{k(i)}) = |\gamma|.$$

Thus $d_\gamma(\alpha_{k(i)} q) = |\gamma|$ for any $i$. $\square$

The following result was established in [Pr1942].

**4.20 Theorem (Preissmann 1942)** *If $M$ is a compact connected Riemannian manifold with* $\sec < 0$*, then every non-trivial abelian subgroup of $\pi_1(M)$ is isomorphic to $\mathbb{Z}$.*

This shows for example that the torus $T^m = \mathbb{R}^m / \mathbb{Z}^m$ ($m \geq 2$) cannot carry a metric with $\sec < 0$.

*Proof*: The universal Riemannian covering $\tilde{M}$ of $M$ is a Hadamard manifold with $\sec < 0$ (note that $\tilde{M}$ is complete since $M$ is), and $M$ is isometric to $\tilde{M}/\Gamma$ for the group $\Gamma \subset \mathrm{Isom}(\tilde{M})$ of deck transformations, which is isomorphic to $\pi_1(M)$. The group $\Gamma$ acts freely and properly discontinuously as well as cocompactly on $\tilde{M}$, because $M$ is compact. By Proposition 4.19, every $\gamma \in \Gamma \setminus \{\mathrm{id}\}$ is axial. Since $\sec < 0$, it follows from Proposition 4.14 that every such $\gamma$ has a unique axis $L_\gamma \subset \tilde{M}$. Furthermore, since $\gamma$ has no fixed point, every line $L \subset \tilde{M}$ preserved by $\gamma$ is an axis and hence equal to $L_\gamma$. Now if $\beta, \gamma \in \Gamma \setminus \{\mathrm{id}\}$ are two commuting elements, then $\gamma(\beta L_\gamma) = \beta(\gamma L_\gamma) = \beta L_\gamma$, thus $\beta L_\gamma = L_\gamma$ and so $L_\beta = L_\gamma$. We conclude that every non-trivial abelian subgroup $A \subset \Gamma$ acts by translation on a line $L \subset \tilde{M}$, and it follows readily that $A$ is isomorphic to $\mathbb{Z}$. □

# Chapter 5

# Triangle comparison

## Some model space geometry

Let $M$ be a metric space with metric $d$. By a *segment* connecting two points $p, q$ in $M$ we mean the image of an isometric embedding $[0, d(p, q)] \to M$ that maps 0 to $p$ and $d(p, q)$ to $q$ (a minimizing geodesic from $p$ to $q$). We will write $pq$ for some such segment (assuming there is one), despite the fact that it need not be uniquely determined by $p$ and $q$. We will use the symbol $|pq|$ as a shorthand for $d(p, q)$, regardless of the existence of a segment $pq$. The *perimeter* of a triple $(p, x, y)$ of points in $M$ is the number

$$\operatorname{Per}(p, x, y) := |px| + |py| + |xy|.$$

By a *hinge* $H_p(x, y)$ in $M$ we mean a collection of three points $p, x, y$ and two non-degenerate segments $px, py$ in $M$; thus $p \notin \{x, y\}$ (but possibly $x = y$). We call $p$ the *vertex*, $x, y$ the *endpoints*, and $px, py$ the *sides* of the hinge. By the perimeter $\operatorname{Per}(H)$ of a hinge $H = H_p(x, y)$ we mean the perimeter of the triple $(p, x, y)$.

Let again $\mathbb{M}^m_\kappa$ denote the $m$-dimensional, complete and simply connected model space of constant sectional curvature $\kappa \in \mathbb{R}$, and recall that

$$D_\kappa := \operatorname{Diam}(\mathbb{M}^m_\kappa) = \begin{cases} \frac{\pi}{\sqrt{\kappa}} & \text{if } \kappa > 0, \\ \infty & \text{if } \kappa \le 0. \end{cases}$$

For $\kappa \in \mathbb{R}$ we denote by $\operatorname{sn}_\kappa \colon \mathbb{R} \to \mathbb{R}$ and $\operatorname{cs}_\kappa \colon \mathbb{R} \to \mathbb{R}$ the solutions of the differential equation $f'' + \kappa f = 0$ satisfying the initial conditions

$$\operatorname{sn}_\kappa(0) = 0, \quad \operatorname{sn}'_\kappa(0) = 1, \qquad \operatorname{cs}_\kappa(0) = 1, \quad \operatorname{cs}'_\kappa(0) = 0.$$

(Recall that if $c \colon \mathbb{R} \to \mathbb{M}^m_\kappa$ is a unit speed geodesic, and if $X$ is a parallel normal vector field along $c$ and $f \colon \mathbb{R} \to \mathbb{R}$ is a smooth function, then $fX$ is a Jacobi field along $c$ if and only if $f'' + \kappa f = 0$.) Explicitly,

$$\mathrm{sn}_\kappa(x) = \sum_{n=0}^{\infty} \frac{(-\kappa)^n}{(2n+1)!} x^{2n+1} = \begin{cases} \frac{1}{\sqrt{\kappa}} \sin(\sqrt{\kappa} x) & \text{if } \kappa > 0, \\ x & \text{if } \kappa = 0, \\ \frac{1}{\sqrt{-\kappa}} \sinh(\sqrt{-\kappa} x) & \text{if } \kappa < 0, \end{cases}$$

$$\mathrm{cs}_\kappa(x) = \sum_{n=0}^{\infty} \frac{(-\kappa)^n}{(2n)!} x^{2n} = \begin{cases} \cos(\sqrt{\kappa} x) & \text{if } \kappa > 0, \\ 1 & \text{if } \kappa = 0, \\ \cosh(\sqrt{-\kappa} x) & \text{if } \kappa < 0. \end{cases}$$

Note that $\mathrm{sn}_\kappa$ is positive on $(0, D_\kappa)$ and strictly increasing on $(-\frac{1}{2}D_\kappa, \frac{1}{2}D_\kappa)$. The identity $\mathrm{cs}_\kappa^2 + \kappa \, \mathrm{sn}_\kappa^2 = 1$ holds, and for $x, y \in \mathbb{R}$,

$$\begin{aligned} \mathrm{sn}_\kappa(x+y) &= \mathrm{sn}_\kappa(x)\,\mathrm{cs}_\kappa(y) + \mathrm{cs}_\kappa(x)\,\mathrm{sn}_\kappa(y), \\ \mathrm{cs}_\kappa(x+y) &= \mathrm{cs}_\kappa(x)\,\mathrm{cs}_\kappa(y) - \kappa\,\mathrm{sn}_\kappa(x)\,\mathrm{sn}_\kappa(y). \\ \kappa\,\mathrm{sn}_\kappa^2\Big(\frac{x}{2}\Big) &= \frac{1 - \mathrm{cs}_\kappa(x)}{2}, \\ \mathrm{cs}_\kappa^2\Big(\frac{x}{2}\Big) &= \frac{1 + \mathrm{cs}_\kappa(x)}{2}. \end{aligned}$$

The law of cosines for $\mathbb{M}_\kappa^m$, $\kappa \in \mathbb{R}$, can be expressed in a unified way as follows.

**5.1 Lemma (law of cosines)** *Let $H_p(x, y)$ be a hinge in $\mathbb{M}_\kappa^2$ with angle $\gamma := \measuredangle_p(x, y) \in [0, \pi]$, and put $a := |py|$, $b := |px|$, $c := |xy|$. Then*

$$\begin{aligned} \mathrm{sn}_\kappa^2\Big(\frac{c}{2}\Big) &= \mathrm{sn}_\kappa^2\Big(\frac{a+b}{2}\Big) - \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b)\cos^2\Big(\frac{\gamma}{2}\Big) \\ &= \mathrm{sn}_\kappa^2\Big(\frac{a-b}{2}\Big) + \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b)\sin^2\Big(\frac{\gamma}{2}\Big). \end{aligned}$$

Note that $\cos^2\big(\frac{\gamma}{2}\big) = \frac{1}{2}(1 + \cos(\gamma))$ and $\sin^2\big(\frac{\gamma}{2}\big) = \frac{1}{2}(1 - \cos(\gamma))$. Multiplying the above formula by $\kappa$ one obtains the (more standard) expression

$$\mathrm{cs}_\kappa(c) = \mathrm{cs}_\kappa(a)\,\mathrm{cs}_\kappa(b) + \kappa\,\mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b)\cos(\gamma)$$

for $\kappa \neq 0$.

*Proof*: Exercise. □

**5.2 Lemma** *Let $\kappa \in \mathbb{R}$ and $a, b \in (0, D_\kappa)$ be fixed. For $\gamma \in [0, \pi]$, let $H_p(x, y)$ be a hinge in $\mathbb{M}_\kappa^2$ such that $|px| = b$, $|py| = a$, and $\measuredangle_p(x, y) = \gamma$, and put $c_{a,b}(\gamma) := d(x, y)$. The function $c_{a,b}$ so defined is continuous and strictly increasing on $[0, \pi]$.*

*Proof*: This follows directly from Lemma 5.1. □

The next lemma goes back to Alexandrov [Al1955] (compare Lemma 4.3.3 in [BuBI2001]).

**5.3 Lemma (Alexandrov)** *Suppose that $H_p(q,y)$ and $H_q(x,y)$ are two hinges in $\mathbb{M}^2_\kappa$ with $|py|, |qy|, |pq| + |qx| < D_\kappa$, and $H_{\bar p}(\bar x, \bar y)$ is a hinge in $\mathbb{M}^2_\kappa$ such that $|\bar p \bar x| = |pq| + |qx|$, $|\bar p \bar y| = |py|$, and $|\bar x \bar y| = |xy|$. Then*

$$\begin{aligned} \measuredangle_q(p,y) + \measuredangle_q(x,y) \le \pi \quad &\Longleftrightarrow \quad \measuredangle_p(q,y) \ge \measuredangle_{\bar p}(\bar x, \bar y), \\ \measuredangle_q(p,y) + \measuredangle_q(x,y) \ge \pi \quad &\Longleftrightarrow \quad \measuredangle_p(q,y) \le \measuredangle_{\bar p}(\bar x, \bar y). \end{aligned}$$

*Proof*: Prolongate $pq$ to a segment $px'$ of length $|px'| = |pq| + |qx|$. Consider the following obvious identities:

$$\pi - \measuredangle_q(p,y) - \measuredangle_q(x,y) = \measuredangle_q(x',y) - \measuredangle_q(x,y), \tag{1}$$

$$|x'y| - |xy| = |x'y| - |\bar x \bar y|, \tag{2}$$

$$\measuredangle_p(x',y) - \measuredangle_{\bar p}(\bar x, \bar y) = \measuredangle_p(q,y) - \measuredangle_{\bar p}(\bar x, \bar y). \tag{3}$$

By Lemma 5.2, the right side of (1) and the left side of (2) have the same sign ($\in \{-1, 0, 1\}$), and also the right side of (2) and the left side of (3) have equal sign. Hence, the same holds for the left side of (1) and the right side of (3). □

## Alexandrov comparisons

Let again $M$ be a metric space, and let $\kappa \in \mathbb{R}$. Given $p, x, y \in M$, a triple $(\bar p, \bar x, \bar y)$ of points in $\mathbb{M}^2_\kappa$ is called a *comparison triple* for $(p,x,y)$ if $|\bar p \bar x| = |px|$, $|\bar p \bar y| = |py|$, and $|\bar x \bar y| = |xy|$.

**5.4 Remark** If $\kappa \le 0$, such a comparison triple always exists, and if $\kappa > 0$, a comparison triple exists if and only if $\mathrm{Per}(p,x,y) \le 2D_\kappa$. This is obvious if one of the distances $a := |py|$, $b := |px|$, and $c := |xy|$ is zero or equal to $D_\kappa$. On the other hand, if $a, b, c \in (0, D_\kappa)$, then it follows from Lemma 5.2 that the function $c_{a,b}$ maps $[0,\pi]$ bijectively onto $I := [|a-b|, a+b]$ or $I := [|a-b|, 2D_\kappa - a - b]$, depending on whether $a + b < D_\kappa$ or $a + b \ge D_\kappa$. In either case, the given number $c$ is contained in $I$, so there exists a unique $\gamma \in [0,\pi]$ such that $c_{a,b}(\gamma) = c$.

**5.5 Definition** Consider a triple $(p,x,y)$ of points in $M$ such that $p \notin \{x,y\}$. In the case $\kappa > 0$, suppose that $|px|, |py| < D_\kappa$ and $\mathrm{Per}(p,x,y) \le 2D_\kappa$. Then any comparison triple $(\bar p, \bar x, \bar y)$ in $\mathbb{M}^2_\kappa$ uniquely determines a hinge $H_{\bar p}(\bar x, \bar y)$ and one defines the *comparison angle* $\measuredangle^\kappa_p(x,y) \in [0,\pi]$ as the hinge angle, thus

$$\measuredangle^\kappa_p(x,y) := \measuredangle_{\bar p}(\bar x, \bar y).$$

For an arbitrary hinge $H_p(x,y)$ in $M$, the (Alexandrov) *angle* or *upper angle* of $H_p(x,y)$ is then defined by

$$\measuredangle_p(x,y) := \limsup_{\substack{u \in px,\, v \in py \\ u,v \to p}} \measuredangle_p^\kappa(u,v).$$

**5.6 Remark** It is not difficult to see that the upper angle $\measuredangle_p(x,y)$ does not depend on the choice of $\kappa \in \mathbb{R}$. Furthermore, if $px, py, pz$ are three non-degenerate segments in $M$, the triangle inequality

$$\measuredangle_p(x,y) + \measuredangle_p(y,z) \geq \measuredangle_p(x,z)$$

holds, see [Al1951] or Proposition I.1.14 in [BrH1999].

**5.7 Definition** Let again $H_p(x,y)$ be a hinge in $M$, and suppose that $\mathrm{Per}(p,x,y) < 2D_\kappa$. Let $(\bar{p},\bar{x},\bar{y})$ be a comparison triple in $\mathbb{M}^2_\kappa$ for $(p,x,y)$, and let $H_{\hat{p}}(\hat{x},\hat{y})$ be a *comparison hinge* in $\mathbb{M}^2_\kappa$ for $H_p(x,y)$; that is, $|\hat{p}\hat{x}| = |px|$, $|\hat{p}\hat{y}| = |py|$, and $\measuredangle_{\hat{p}}(\hat{x},\hat{y}) = \measuredangle_p(x,y)$. We are interested in the following *comparison properties for curvature* $\geq \kappa$ that $H_p(x,y)$ may or may not have:

($\mathrm{A}_\kappa$) (Angle comparison) $\measuredangle_p(x,y) \geq \measuredangle_{\bar{p}}(\bar{x},\bar{y})$ $(= \measuredangle_p^\kappa(x,y))$;

($\mathrm{H}_\kappa$) (Hinge comparison) $|xy| \leq |\hat{x}\hat{y}|$;

($\mathrm{D}_\kappa$) (Distance comparison) $|uv| \geq |\bar{u}\bar{v}|$ whenever $u \in px$, $v \in py$, $\bar{u} \in \bar{p}\bar{x}$, $\bar{v} \in \bar{p}\bar{y}$, and $|pu| = |\bar{p}\bar{u}|$, $|pv| = |\bar{p}\bar{v}|$.

The corresponding *comparison properties for curvature* $\leq \kappa$, denoted ($\mathrm{A}^\kappa$), ($\mathrm{H}^\kappa$), and ($\mathrm{D}^\kappa$), are defined analogously, just with reversed inequalities.

**5.8 Lemma** *For an individual hinge $H_p(x,y)$ in $M$ with* $\mathrm{Per}(p,x,y) < 2D_\kappa$*,*

$$(\mathrm{D}_\kappa) \Rightarrow (\mathrm{A}_\kappa) \Leftrightarrow (\mathrm{H}_\kappa) \quad \textit{and} \quad (\mathrm{D}^\kappa) \Rightarrow (\mathrm{A}^\kappa) \Leftrightarrow (\mathrm{H}^\kappa).$$

For the implications $(\mathrm{A}_\kappa) \Rightarrow (\mathrm{D}_\kappa)$ and $(\mathrm{A}^\kappa) \Rightarrow (\mathrm{D}^\kappa)$, see Lemma 5.9 below.

*Proof*: Suppose that $H_p(x,y)$ satisfies ($\mathrm{D}_\kappa$). With the above notation, it follows from the inequality $|uv| \geq |\bar{u}\bar{v}|$ and Lemma 5.2 that $\measuredangle_p^\kappa(u,v) \geq \measuredangle_{\bar{p}}(\bar{u},\bar{v}) = \measuredangle_{\bar{p}}(\bar{x},\bar{y})$. Taking the upper limit for $u,v \to p$, we conclude that ($\mathrm{A}_\kappa$) holds. Lemma 5.2 also shows that $\measuredangle_p(x,y) = \measuredangle_{\hat{p}}(\hat{x},\hat{y}) \geq \measuredangle_{\bar{p}}(\bar{x},\bar{y})$ if and only if $|\hat{x}\hat{y}| \geq |\bar{x}\bar{y}| = |xy|$; that is, $(\mathrm{A}_\kappa) \Leftrightarrow (\mathrm{H}_\kappa)$. The second part of the lemma is shown analogously. □

We call a segment $px$ in a metric space *balanced* if, for every non-degenerate segment $qy$ with $q \in px \setminus \{p,x\}$, the (upper) angles formed by $qy$ and the subsegments $qp, qx$ of $px$ satisfy $\measuredangle_q(p,y) + \measuredangle_q(x,y) = \pi$. Note that, by the triangle inequality for angles (Remark 5.6), the inequality $\measuredangle_q(p,y) + \measuredangle_q(x,y) \geq \pi$ always holds, since $\measuredangle_q(p,x) = \pi$. Obviously, in a Riemannian manifold every segment is balanced.

**5.9 Lemma** *Let $H_p(x, y)$ be a hinge in $M$ with* $\mathrm{Per}(p, x, y) < 2D_\kappa$*, and suppose that for every pair of points in $px \cup py$ there is a connecting segment in $M$. If every hinge with one side contained in $px$ or $py$ and the opposite endpoint on $py$ or $px$, respectively, satisfies* ($\mathrm{A}^\kappa$)*, then $H_p(x, y)$ satisfies* ($\mathrm{D}^\kappa$)*. The analogous result for* ($\mathrm{A}_\kappa$) *and* ($\mathrm{D}_\kappa$) *holds if the segments $px, py$ are balanced.*

*Proof*: Let $(\bar{p}, \bar{x}, \bar{y})$ be a comparison triple in $\mathbb{M}^2_\kappa$ for $(p, x, y)$, and let $u, v \neq p$ and $\bar{u}, \bar{v}$ be given as in ($\mathrm{D}^\kappa$). First we show that $|uy| \leq |\bar{u}\bar{y}|$. Omitting some trivial cases, we assume that $u \neq x, y$. The two hinges formed by a segment $uy$ and one of the subsegments $up, ux$ of $px$ satisfy ($\mathrm{A}^\kappa$) by assumption, thus

$$\measuredangle^\kappa_u(p, y) + \measuredangle^\kappa_u(x, y) \geq \measuredangle_u(p, y) + \measuredangle_u(x, y) \geq \pi. \tag{1}$$

Lemma 5.3 then shows that $\measuredangle^\kappa_p(u, y) \leq \measuredangle_{\bar{p}}(\bar{x}, \bar{y}) = \measuredangle_{\bar{p}}(\bar{u}, \bar{y})$, and Lemma 5.2 yields $|uy| \leq |\bar{u}\bar{y}|$. Now an analogous argument shows that $|uv| \leq |\tilde{u}\tilde{v}|$ if $(\tilde{p}, \tilde{u}, \tilde{y})$ is a comparison triple for $(p, u, y)$ and $\tilde{v} \in \tilde{p}\tilde{y}$ is such that $|pv| = |\tilde{p}\tilde{v}|$. Since $|\tilde{u}\tilde{y}| = |uy| \leq |\bar{u}\bar{y}|$, we get that $\measuredangle_{\tilde{p}}(\tilde{u}, \tilde{v}) = \measuredangle_{\tilde{p}}(\tilde{u}, \tilde{y}) \leq \measuredangle_{\bar{p}}(\bar{u}, \bar{y}) = \measuredangle_{\bar{p}}(\bar{u}, \bar{v})$ and hence $|\tilde{u}\tilde{v}| \leq |\bar{u}\bar{v}|$ by Lemma 5.2. Thus $|uv| \leq |\bar{u}\bar{v}|$, as required.

The corresponding result for ($\mathrm{A}_\kappa$) and ($\mathrm{D}_\kappa$) is shown in exactly the same way, with all inequalities reversed, except that the second relation in (1) is turned into an equality, holding by assumption. □

**5.10 Definition** A metric space $M$ is called a *space of curvature* $\geq \kappa$ or $\leq \kappa$ *in the sense of Alexandrov* if every point $q$ has a neighborhood $U_q$ such that any two points in $U_q$ are connected by a segment in $M$ and every hinge $H_p(x, y)$ with $p, x, y \in U_q$ and $\mathrm{Per}(p, x, y) < 2D_\kappa$ satisfies ($\mathrm{D}_\kappa$) or ($\mathrm{D}^\kappa$), respectively.

Again due to Lemma 5.2, the upper angle between two segments in such a space $M$ always exists as a limit, by monotonicity.

**5.11 Lemma** *In an Alexandrov space of curvature $\geq \kappa \in \mathbb{R}$, every segment is balanced.*

*Proof*: Let $px, qy$ be two non-degenerate segments in $M$ such that $q \in px \setminus \{p, x\}$. Let $u \in qp$, $v \in qx$, $w \in qy$ be points distinct from $q$, and assume that $u \neq w$. If $u, v, w$ are sufficiently close to $q$, then there is a segment $uw$ such that the hinge $H_u(v, w)$ with $uv \subset px$ satisfies ($\mathrm{D}_\kappa$). Let $(\bar{u}, \bar{v}, \bar{w})$ be a comparison triple in $\mathbb{M}^2_\kappa$ for $(u, v, w)$, and let $\bar{q} \in \bar{u}\bar{v}$ be the point with $|\bar{q}\bar{u}| = |qu|$. Then $|qw| \geq |\bar{q}\bar{w}|$ and so $\measuredangle^\kappa_u(q, w) \geq \measuredangle_{\bar{u}}(\bar{q}, \bar{w}) = \measuredangle_{\bar{u}}(\bar{v}, \bar{w})$ by Lemma 5.2. Now Lemma 5.3 shows that $\measuredangle^\kappa_q(u, w) + \measuredangle^\kappa_q(v, w) \leq \pi$. Letting $u, v, w$ tend to $q$ we get that $\measuredangle_q(p, y) + \measuredangle_q(x, y) \leq \pi$. □

**5.12 Theorem** *A connected Riemannian manifold $M$ is a space of curvature $\geq \kappa$ or $\leq \kappa \in \mathbb{R}$ in the sense of Alexandrov if and only if* $\sec \geq \kappa$ *or* $\sec \leq \kappa$*, respectively.*

*Proof*: Suppose that $\sec \geq \kappa$ or $\sec \leq \kappa$. If $p \in M$, $r \in (0, D_\kappa]$, and $\exp_p : B_r \to B(p,r)$ is a diffeomorphism, then it follows easily from Corollary 3.19 that every hinge $H_p(x,y)$ in $M$ with $x, y \in B(p, r/2)$ satisfies ($\mathrm{H}_\kappa$) or ($\mathrm{H}^\kappa$), respectively (note that $d(x,y) < r$ and every curve of length $< r$ from $x$ to $y$ lies in $B(p,r)$). Now it follows from (the proof of) Corollary 1.22, Lemma 5.8, Lemma 5.9 and the fact that segments in $M$ are balanced that $M$ has curvature $\geq \kappa$ or $\leq \kappa$ in the sense of Alexandrov.

The proof of the other implication is left as an exercise. $\square$

## Toponogov's Theorem

**5.13 Lemma** *Let $\kappa \in \mathbb{R}$, let $M$ be a metric space, and let $H_p(x,y)$ be a hinge in $M$ with* $\mathrm{Per}(p,x,y) < 2D_\kappa$. *Suppose that there exist a point $q \in px \setminus \{p,x,y\}$ and a segment $qy$ such that each of the three hinges $H_p(q,y), H_q(p,y), H_q(x,y)$ with sides in $px \cup py \cup qy$ satisfies* ($\mathrm{A}_\kappa$)*, and $\measuredangle_q(p,y) + \measuredangle_q(x,y) = \pi$. Then $H$ satisfies* ($\mathrm{A}_\kappa$) *as well.*

*Proof*: Note that $\mathrm{Per}(p,q,y), \mathrm{Per}(q,x,y) \leq \mathrm{Per}(p,x,y) < 2D_\kappa$. Since $H_p(q,y)$ satisfies ($\mathrm{A}_\kappa$), $\measuredangle_p(x,y) = \measuredangle_p(q,y) \geq \measuredangle^\kappa_p(q,y)$. By the remaining assumptions,

$$\measuredangle^\kappa_q(p,y) + \measuredangle^\kappa_q(x,y) \leq \measuredangle_q(p,y) + \measuredangle_q(x,y) = \pi,$$

hence Lemma 5.3 shows that $\measuredangle^\kappa_p(q,y) \geq \measuredangle^\kappa_p(x,y)$. Thus $\measuredangle_p(x,y) \geq \measuredangle^\kappa_p(x,y)$. $\square$

**5.14 Proposition** *Let $\kappa \in \mathbb{R}$, and let $M$ be a metric space such that every segment in $M$ is balanced and every pair of points in $M$ at distance $< D_\kappa$ is connected by a segment. Let $H_p(x,y)$ be a hinge in $M$ with* $\mathrm{Per}(p,x,y) < 2D_\kappa$. *If every hinge $H_{p'}(x',y')$ in $M$ with* $\mathrm{Per}(p',x',y') < \frac{4}{5}\,\mathrm{Per}(p,x,y)$ *and an endpoint on $px \cup py$ satisfies* ($\mathrm{A}_\kappa$)*, then $H_p(x,y)$ satisfies* ($\mathrm{A}_\kappa$) *as well.*

*Proof*: Part I. First we prove that *if $H_0 = H_{p_0}(x_0,y_0)$ is a hinge in $M$ with sides of length $a := |p_0y_0|$ and $b := |p_0x_0| < \frac{1}{5}a$, where $a + b < D_\kappa$, and if every hinge $H_{p'}(x',y')$ in $M$ with* $\mathrm{Per}(p',x',y') < \frac{4}{5}\,\mathrm{Per}(p_0,x_0,y_0)$ *and $\{x',y'\} \cap \{x_0,y_0\} \neq \emptyset$ satisfies* ($\mathrm{A}_\kappa$)*, then $H_0$ satisfies* ($\mathrm{A}_\kappa$) *as well.*

Starting from $H_0$, we will inductively construct a particular sequence of hinges $H_n = H_{p_n}(x_n,y_n)$ in $M$ such that $\{x_n,y_n\} = \{x_0,y_0\}$ and the numbers $l_n := |p_nx_n| + |p_ny_n|$ satisfy

$$a + b = l_0 \geq l_1 \geq l_2 \geq \ldots \geq |x_0y_0|; \tag{1}$$

furthermore, for $n \geq 1$, $|p_nx_n| = b' := \frac{2}{5}a$ and hence

$$|p_ny_n| \geq |x_ny_n| - |p_nx_n| = |x_0y_0| - b' \geq a - b - b' > b'. \tag{2}$$

The hinge $H_0$ is already given. Let $\gamma_0$ denote its angle. For $n \geq 1$, if $H_{n-1}$ is constructed, let $p_n \in p_{n-1}y_{n-1}$ be the point at distance $b'$ from $y_{n-1}$, and put $x_n := y_{n-1}$ and $y_n := x_{n-1}$. The sides of $H_n$ are the subsegment $p_nx_n$ of $p_{n-1}y_{n-1}$ and an arbitrary segment $p_ny_n$. Let $\gamma_n$ denote the angle of $H_n$, and note that the adjacent angle between $p_ny_n$ and the segment $p_np_{n-1} \subset p_{n-1}y_{n-1}$ equals $\pi - \gamma_n$, because segments in $M$ are balanced. Clearly (1) holds. Note also that

$$\text{(3)} \qquad \operatorname{Per}(p_{n-1}, p_n, y_n) \leq 2(l_{n-1} - b') \leq 2(a + b - b') < \tfrac{8}{5}a \leq \tfrac{4}{5}\operatorname{Per}(p_0, x_0, y_0).$$

Now we will construct a sequence of hinges $\bar{H}_n := H_{\bar{p}_n}(\bar{x}_n, \bar{y}_n)$ in $\mathbb{M}^2_\kappa$ such that $|\bar{p}_n\bar{x}_n| = |p_nx_n|$, $|\bar{p}_n\bar{y}_n| = |p_ny_n|$,

$$\text{(4)} \qquad |\bar{x}_0\bar{y}_0| \geq |\bar{x}_1\bar{y}_1| \geq |\bar{x}_2\bar{y}_2| \geq \ldots,$$

and such that the angle $\bar{\gamma}_n$ of $\bar{H}_n$ is greater than or equal to $\gamma_n$. Let $\bar{H}_0$ be a comparison hinge for $H_0$, thus $\bar{\gamma}_0 = \gamma_0$. For $n \geq 1$, given $\bar{H}_{n-1}$, let $\bar{p}_n \in \bar{p}_{n-1}\bar{y}_{n-1}$ be the point at distance $b'$ from $\bar{y}_{n-1}$, put $\bar{x}_n := \bar{y}_{n-1}$, and choose $\bar{y}_n$ such that $(\bar{p}_{n-1}, \bar{p}_n, \bar{y}_n)$ is a comparison triple for $(p_{n-1}, p_n, y_n)$. This determines $\bar{H}_n$. Put $\bar{\omega}_n := \measuredangle_{\bar{p}_{n-1}}(\bar{p}_n, \bar{y}_n) = \measuredangle_{\bar{p}_{n-1}}(\bar{x}_n, \bar{y}_n)$. In view of (3), and since $y_n \in \{x_0, y_0\}$, the inequalities $\gamma_{n-1} \geq \bar{\omega}_n$ and $\pi - \gamma_n \geq \pi - \bar{\gamma}_n$ hold by assumption. Hence, $\bar{\gamma}_{n-1} \geq \gamma_{n-1} \geq \bar{\omega}_n$ and so $|\bar{x}_{n-1}\bar{y}_{n-1}| \geq |\bar{x}_n\bar{y}_n|$ by Lemma 5.2.

Now, if $n \to \infty$, then

$$|\bar{p}_{n-1}\bar{p}_n| + |\bar{p}_{n-1}\bar{y}_n| - |\bar{p}_n\bar{y}_n| = l_{n-1} - l_n \to 0$$

by (1), consequently $\bar{\omega}_n \to \pi$ and $\bar{\gamma}_n \to \pi$ (note that $|\bar{p}_{n-1}\bar{p}_n| = |p_{n-1}y_{n-1}| - b' \geq a - b - 2b' > 0$ by (2), and $|\bar{p}_{n-1}\bar{y}_n| = b' > 0$). This implies in turn that

$$l_n - |\bar{x}_n\bar{y}_n| = |\bar{p}_n\bar{x}_n| + |\bar{p}_n\bar{y}_n| - |\bar{x}_n\bar{y}_n| \to 0$$

as $n \to \infty$ (recall that $l_n \leq a + b < D_\kappa$). In view of (1) and (4), this gives $|\bar{x}_0\bar{y}_0| \geq |x_0y_0|$, so $H_0$ satisfies ($\mathrm{H}_\kappa$) and hence also ($\mathrm{A}_\kappa$).

Part II. Let a hinge $H_p(x, y)$ with $\operatorname{Per}(p, x, y) < 2D_\kappa$ be given. If $y \in px$, then $\measuredangle^\kappa_p(x, y) = 0$ and ($\mathrm{A}_\kappa$) holds trivially. Suppose now that $a := d(y, px) > 0$. Note that for all $x' \in px$, $|x'y| \leq \frac{1}{2}\operatorname{Per}(p, x, y) < D_\kappa$. Choose $b \in (0, \frac{1}{5}a)$ such that $|x'y| + b < D_\kappa$ for all $x' \in px$. Now subdivide $px$ into finitely many subsegments of length at most $b$. It follows from the result of Part I that every hinge $H_{x'}(x'', y)$ with $x'x'' \subset px$ and $|x'x''| \leq b$ satisfies ($\mathrm{A}_\kappa$). Now use Lemma 5.13 repeatedly to show that $H_p(x, y)$ satisfies ($\mathrm{A}_\kappa$) as well. □

**5.15 Theorem (Toponogov)** *Let $\kappa \in \mathbb{R}$, and let $M$ be a complete metric space of curvature $\geq \kappa$ in the sense of Alexandrov. Suppose that every pair of points in $M$ at distance $< D_\kappa$ is connected by a segment. Then every hinge $H_p(x, y)$ in $M$ with* $\operatorname{Per}(p, x, y) < 2D_\kappa$ *satisfies* ($\mathrm{A}_\kappa$), ($\mathrm{H}_\kappa$), *and* ($\mathrm{D}_\kappa$).

Note that this applies, of course, to complete Riemannian manifolds with sectional curvature greater than or equal to $\kappa$. For this case, the result was established in [To1959] (see also [To1957] and [To1958]). Purely metric proofs were given later in [Pl1991], [BurGP1992], and [Pl1996].

A closer look at the proof (of Lemma 5.13 and Part II of Proposition 5.14) reveals that the comparisons ($\mathrm{A}_\kappa$) and ($\mathrm{H}_\kappa$) hold also for *generalized hinges* where one of the sides, $px$ say, is just a (locally minimizing) geodesic segment of length $L(px) \leq |py| + |xy|$. The condition $\mathrm{Per}(p, x, y) < 2D_\kappa$ is then replaced by the assumption that $L(px) + |py| + |xy| < 2D_\kappa$, and comparison triples or hinges in $\mathbb{M}^2_\kappa$ are chosen such that $|\bar{p}\bar{x}| = L(px)$.

*Proof*: Recall that by Lemma 5.11 all segments in $M$ are balanced. By Lemma 5.8 and Lemma 5.9, it thus suffices to prove that every hinge in $M$ with perimeter less than $2D_\kappa$ satisfies ($\mathrm{A}_\kappa$). Suppose to the contrary that there exists a hinge $H$ in $M$ with $\mathrm{Per}(H) < 2D_\kappa$ that does not satisfy ($\mathrm{A}_\kappa$). Then, by Proposition 5.14, there exists a hinge $H_1$ with $\mathrm{Per}(H_1) < \frac{4}{5}\,\mathrm{Per}(H)$ and an endpoint on the union of the sides of $H$ such that $H_1$ does not satisfy ($\mathrm{A}_\kappa$) either. Inductively, for $n = 2, 3, \ldots$, there exist hinges $H_n$ such that $\mathrm{Per}(H_n) < \frac{4}{5}\,\mathrm{Per}(H_{n-1}) < \left(\frac{4}{5}\right)^n \mathrm{Per}(H)$, some endpoint of $H_n$ lies on the union of the sides of $H_{n-1}$, and $H_n$ does not satisfy ($\mathrm{A}_\kappa$). Let $p_n$ denote the vertex of $H_n$. Clearly the sequence $(p_n)$ is Cauchy and thus converges to a point $q \in M$. However, since $M$ has curvature $\geq \kappa$, all hinges with vertex and endpoints in an appropriate neighborhood of $q$ satisfy ($\mathrm{A}_\kappa$). This gives a contradiction, as $p_n \to q$ and $\mathrm{Per}(H_n) \to 0$. □

**5.16 Theorem** *Let $M$ be a complete and geodesic Alexandrov space of curvature $\geq \kappa > 0$. Suppose that for every segment $px$ in $M$ with $|px| > D_\kappa$ and midpoint $q$ there exists a non-trivial segment $qy$ perpendicular to $px$* (*this excludes* 1-*dimensional spaces of diameter $> D_\kappa$*). *Then every triple of points $p, x, y$ in $M$ has* $\mathrm{Per}(p, x, y) \leq 2D_\kappa$; *in particular* $\mathrm{Diam}(M) \leq D_\kappa$. *If* $\mathrm{Per}(p, x, y) = 2D_\kappa$ *and $|px|, |py| \in (0, D_\kappa)$, then there are unique segments $px$ and $py$, and $\measuredangle_p(x, y) = \pi$.*

*Proof*: First we show that $\mathrm{Diam}(M) \leq D_\kappa$. Suppose that $px$ is a segment in $M$ of length $|px| \in (D_\kappa, 2D_\kappa)$. Let $q$ be its midpoint, and let $qy$ be a non-trivial segment perpendicular to $px$ such that $\mathrm{Per}(p, q, y), \mathrm{Per}(q, x, y) < 2D_\kappa$. Consider two segments $\bar{p}\bar{q}$ and $\bar{q}\bar{x}$ in $\mathbb{M}^2_\kappa$ such that $|\bar{p}\bar{q}| = |\bar{q}\bar{x}| = \frac{1}{2}|px|$ and $\measuredangle_{\bar{q}}(\bar{p}, \bar{x}) = \pi$, and let $\bar{q}\bar{y}$ be a perpendicular segment of length $|qy|$. Since $\bar{p}$ and $\bar{x}$ are not antipodal (and $\bar{p} \neq \bar{x}$), $\bar{p}, \bar{y}, \bar{x}$ do not lie on a great circle, hence $\mathrm{Per}(\bar{p}, \bar{y}, \bar{x}) < 2D_\kappa = \mathrm{Per}(\bar{p}, \bar{q}, \bar{x})$ and so

$$|\bar{p}\bar{y}| + |\bar{y}\bar{x}| < |\bar{p}\bar{q}| + |\bar{q}\bar{x}| = |px|.$$

By Theorem 5.15 (hinge comparison ($\mathrm{H}_\kappa$)), the left side is greater than or equal to $|py| + |yx|$. Thus $|py| + |yx| < |px|$, in contradiction to the triangle inequality. Since $M$ is geodesic, it follows that $\mathrm{Diam}(M) \leq D_\kappa$.

Next, suppose that there exists a hinge $H_p(x, y)$ in $M$ with $P := \mathrm{Per}(p, x, y) > 2D_\kappa$. Note that $|px|, |py|, |xy| \le D_\kappa < \frac{1}{2}P$; in particular there is a point $q \in px \setminus \{p, x\}$ such that $|pq| + |py| = \frac{1}{2}P$. Let $\kappa' \in (0, \kappa)$ be such that $P = 2D_{\kappa'}$, and let $(\bar{p}, \bar{x}, \bar{y})$ be a comparison triple in $\mathbb{M}^2_{\kappa'}$ for $(p, x, y)$. Then the point $\bar{q} \in \bar{p}\bar{x}$ with $|\bar{p}\bar{q}| = |pq|$ is antipodal to $\bar{y}$, thus $|\bar{q}\bar{y}| = \frac{1}{2}P$. Now it follows from distance comparison ($\mathrm{D}_{\kappa''}$) for $\kappa'' < \kappa'$ and a limit argument that $|qy| \ge \frac{1}{2}P > D_\kappa$, in contradiction to $\mathrm{Diam}(M) \le D_\kappa$. This shows that all triples in $M$ have perimeter at most $2D_\kappa$.

Finally, suppose that $P := \mathrm{Per}(p, x, y) = 2D_\kappa$ and $|px|, |py| \in (0, D_\kappa)$. Note that $\measuredangle^\kappa_p(x, y) = \pi$. Hence, by angle comparison ($\mathrm{A}_{\kappa'}$) for $\kappa' < \kappa$ and a limit argument, $\measuredangle_p(x, y) = \pi$ for any choice of segments $px, py$. It then follows easily that $px, py$ are in fact unique. ◻

## Open manifolds of non-negative curvature

To illustrate the utility of Toponogov's Theorem, we will now discuss some of the results of Cheeger, Gromoll, and Meyer on the global shape of complete non-compact manifolds of non-negative sectional curvature. The general references are [GrM1969], [ChG1972], [Sh1974], and Section 8 in [ChE1975].

We start with a brief general discussion of Busemann functions.

Let $M$ be a complete Riemannian manifold. A *ray* $\varrho$ in $M$ is a geodesic $\varrho\colon [0, \infty) \to M$, parametrized by arc length, such that $d(\varrho(s), \varrho(t)) = |s - t|$ for all $s, t \in [0, \infty)$. If $M$ is non-compact, then for every point $p \in M$ there is at least one ray $\varrho$ with $\varrho(0) = p$. To see this, choose a sequence of minimizing geodesics $\varrho_i\colon [0, d(p, q_i)] \to M$ from $p$ to $q_i$, where $d(p, q_i) \to \infty$. Since $M$ is a proper metric space, it follows easily that some subsequence converges uniformly on compact sets to a ray emanating from $p$.

Now let a ray $\varrho$ in $M$ be given. For $x \in M$ and $t \ge s \ge 0$,

$$d(x, \varrho(t)) - t \le d(x, \varrho(s)) + d(\varrho(s), \varrho(t)) - t = d(x, \varrho(s)) - s,$$

furthermore the function $t \mapsto d(x, \varrho(t)) - t$ is bounded from below by $-d(x, \varrho(0))$. Therefore the limit

$$b_\varrho(x) := \lim_{t \to \infty} d(x, \varrho(t)) - t$$

exists. This defines the *Busemann function* $b_\varrho\colon M \to \mathbb{R}$ of $\varrho$. Note that $b_\varrho(\varrho(s)) = -s$ for all $s \ge 0$, and

$$|b_\varrho(x) - b_\varrho(y)| \le d(x, y)$$

for all $x, y \in M$, that is, $b_\varrho$ is 1-Lipschitz.

For $r \in \mathbb{R}$, the level set $\{b_\varrho = r\}$ is called a *horosphere*, the sublevel set $B(\varrho, r) := \{b_\varrho < r\}$ an (open) *horoball*. Clearly

$$B(\varrho, r) = \bigcup_{t \in [0, \infty)} B(\varrho(t), r + t);$$

indeed, both sides are equal to the set of all $x \in M$ with the property that $d(x, \varrho(t)) - t < r$ for some $t \geq 0$.

**5.17 Lemma** *Let $M$ be a complete Riemannian manifold. Given a ray $\varrho$ and a point $q \in M$, there exists a ray $\sigma$ such that $\sigma(0) = q$ and*

$$b_\varrho(\sigma(s)) = b_\varrho(q) - s$$

*for all $s \geq 0$. The Busemann function $b_\sigma$ of any such ray $\sigma$ satisfies*

$$b_\sigma(x) \geq b_\varrho(x) - b_\varrho(q)$$

*for every $x \in M$.*

*Proof*: For every $t \geq 0$, let $\sigma_t \colon [0, d(q, \varrho(t)] \to M$ be a minimizing geodesic from $q$ to $\varrho(t)$. Then, for a fixed $s \geq 0$ and for all sufficiently large $t$,

$$b_\varrho(q) - s \leq b_\varrho(\sigma_t(s)) \leq b_\varrho(\varrho(t)) + d(\sigma_t(s), \varrho(t)) = -t + d(q, \varrho(t)) - s.$$

As $t \to \infty$, the last term tends to $b_\varrho(q) - s$, hence $b_\varrho(\sigma_t(s)) \to b_\varrho(q) - s$ as well. For some sequence $t_i \to \infty$ the $\sigma_{t_i}$ converge, uniformly on compact sets, to the desired ray $\sigma$.

Since $d(x, \sigma(s)) - s \geq b_\varrho(x) - b_\varrho(\sigma(s)) - s = b_\varrho(x) - b_\varrho(q)$ for all $s \geq 0$, the Busemann function of $\sigma$ satisfies $b_\sigma(x) \geq b_\varrho(x) - b_\varrho(q)$ for all $x \in M$, as claimed. □

Let further $M$ denote a complete Riemannian manifold. A set $C \subset M$ is called *totally convex* if every geodesic $c \colon [0, 1] \to M$ (minimizing or not) with $c(0), c(1) \in C$ satisfies $c(\lambda) \in C$ for all $\lambda \in [0, 1]$.

A function $f \colon C \to \mathbb{R}$ on a totally convex set $C \subset M$ is called *concave* if for every geodesic $c \colon [0, 1] \to C$ and every $\lambda \in [0, 1]$, we have

$$f(c(\lambda)) \geq (1 - \lambda) f(c(0)) + \lambda f(c(1)).$$

Note that then, for every $r \in \mathbb{R}$, the superlevel set $\{f \geq r\}$ is totally convex. To check the concavity of $f$, it clearly suffices to verify the above inequality for every *minimizing* geodesic $c \colon [0, 1] \to C$.

**5.18 Lemma** *Let $M$ be a complete Riemannian manifold of non-negative sectional curvature, and let $\varrho$ be a ray in $M$. Then the Busemann function $b_\varrho \colon M \to \mathbb{R}$ is concave, thus for every $r \in \mathbb{R}$ the horoball complement $C := M \setminus B(\varrho, r)$ is a closed totally convex subset of $M$.*

*Proof*: Let $c\colon [0,1] \to M$ be a minimizing geodesic from $x$ to $y \neq x$. Let $\lambda \in (0,1)$, and put $q := c(\lambda)$. By Lemma 5.17 there exists a ray $\sigma$ with $\sigma(0) = q$ whose Busemann function satisfies

$$b_\sigma(z) \geq b_\varrho(z) - b_\varrho(q)$$

for all $z \in M$. Let $\gamma$ denote the angle between $\sigma'(0)$ and $c'(\lambda)$. Now choose a segment $\bar{x}\bar{y}$ and a ray $\bar{\sigma}$ in $\mathbb{R}^2$ such that $d(\bar{x},\bar{y}) = d(x,y)$, $\bar{\sigma}(0) = \bar{q} := (1-\lambda)\bar{x} + \lambda\bar{y}$, and $\bar{\sigma}$ forms the same angle $\gamma$ with $\bar{q}\bar{y}$. By Theorem 5.15 (hinge comparison ($\mathrm{H}_0$)), $d(x,\sigma(s)) \leq d(\bar{x},\bar{\sigma}(s))$ for all $s \geq 0$, therefore $b_\sigma(x) \leq b_{\bar{\sigma}}(\bar{x})$. Together with the above inequality for $z = x$ this gives $b_\varrho(x) - b_\varrho(q) \leq b_{\bar{\sigma}}(\bar{x})$. Likewise, $b_\varrho(y) - b_\varrho(q) \leq b_{\bar{\sigma}}(\bar{y})$. Combining these two inequalities we obtain

$$(1-\lambda)b_\varrho(x) + \lambda b_\varrho(y) - b_\varrho(q) \leq (1-\lambda)b_{\bar{\sigma}}(\bar{x}) + \lambda b_{\bar{\sigma}}(\bar{y}),$$

and the right side equals zero since $b_{\bar{\sigma}}$ is affine and $b_{\bar{\sigma}}((1-\lambda)\bar{x} + \lambda\bar{y}) = b_{\bar{\sigma}}(\bar{q}) = 0$. □

**5.19 Proposition** *Let $M$ be a complete non-compact Riemannian manifold of non-negative sectional curvature. For every $p \in M$ there exist a number $t \geq 0$ and a family $(C_r)_{r\in(-\infty,t]}$ of compact totally convex subsets $C_r \neq \emptyset$ of $M$ such that $p \in \partial C_0$,*

$$C_s = \{q \in C_r : d(q,\partial C_r) \geq s - r\}$$

*whenever $r < s \leq t$, and $C_t$ has empty interior.*

In particular, for $r < s \leq t$, $C_r$ contains the closed $(s-r)$-neighborhood of $C_s$. Hence the family $(C_r)$ exhausts $M$, that is, $\bigcup_r C_r = M$.

We also note that the curvature assumption will only be used through Lemma 5.18.

*Proof*: Denote by $R$ the set of all rays starting at the given point $p$. For every $r \in \mathbb{R}$, define

$$C_r := \bigcap_{\varrho\in R}(M \setminus B(\varrho,r)) = \{x \in M : b_\varrho(x) \geq r \text{ for all } \varrho \in R\}.$$

Clearly $C_r$ is closed and totally convex (Lemma 5.18). For every ray $\varrho \in R$, $\varrho(0) = p$ is a boundary point of $M \setminus B(\varrho,0)$, so $p \in \partial C_0$. Obviously $C_s \subset C_r$ whenever $r \leq s \in \mathbb{R}$. In particular, for the compactness assertion, it suffices to show that every $C_r$ with $r \leq 0$ is compact. Suppose to the contrary that some such $C_r$ is non-compact. Then there exists a sequence of points $q_i \in C_r$ so that $d(p,q_i) \to \infty$ for $i \to \infty$. Since also $p \in C_r$ and $C_r$ is totally convex, for every $q_i$ there is a minimizing geodesic $\varrho_i\colon [0,d(p,q_i)] \to M$ from $p$ to $q_i$ with image in $C_r$. It follows that there is a ray $\varrho \in R$ with image in $C_r$, in contradiction to the fact that $\varrho((|r|,\infty)) \subset B(\varrho,r)$.

Suppose now that $r < s$ and $q \in C_r$. We want to prove that $q \in C_s$ if and only if $d(q, \partial C_r) \geq s - r$. If $q \in C_s$, note that for every $x \in M \setminus C_r$ there is a ray $\varrho \in R$ such that $b_\varrho(x) < r$ and thus $d(q, x) \geq b_\varrho(q) - b_\varrho(x) > s - r$; consequently $d(q, \partial C_r) = d(q, M \setminus C_r) \geq s - r$. Conversely, assume that $q \notin C_s$. Then $b_\varrho(q) < s$ for some $\varrho \in R$, and by Lemma 5.17 there exists a ray $\sigma$ such that $\sigma(0) = q$ and $b_\varrho(\sigma(s - r)) = b_\varrho(q) - (s - r) < r$. Hence $\sigma(s - r)$ belongs to the open set $M \setminus C_r$ and so $d(q, \partial C_r) < d(q, \sigma(s - r)) = s - r$.

Finally, let $t := \sup\{r \in \mathbb{R} : C_r \neq \emptyset\}$. From the properties of the sets $C_r$ already shown it is easily seen that $0 \leq t < \infty$, $C_r \neq \emptyset$ for all $r \leq t$, and the interior of $C_t$ is empty. $\square$

For a closed and totally convex subset $C \neq \emptyset$ of a complete $m$-dimensional Riemannian manifold $M$, we now define the dimension $\dim(C)$ as the largest integer $l \in \{0, \ldots, m\}$ such that $C$ contains a non-empty $l$-dimensional submanifold of $M$. Then we let $L(C)$ denote the union of all submanifolds $L \subset M$ of dimension $l = \dim(C)$ contained in $C$.

**5.20 Proposition** *Suppose that $M$ is a complete $m$-dimensional Riemannian manifold and $C \neq \emptyset$ is a closed totally convex subset of $M$ with* $\dim(C) = l$. *Then $L(C)$ is an $l$-dimensional submanifold of $M$. If $c\colon [0, 1] \to M$ is a geodesic with $c(0) \in L(C)$ and $c(1) \in C$, then* $c([0, 1)) \subset L(C)$; *in particular $L(C)$ is totally convex and $C$ is the closure of $L(C)$.*

Note also that every totally convex submanifold $N \subset M$ is totally geodesic. On the other hand, a great circle in $S^m \subset \mathbb{R}^{m+1}$ is totally geodesic but not totally convex.

*Proof*: We first show that for every $p \in L(C)$ there exists a neighborhood $V$ of $p$ in $M$ such that $L(C) \cap V = C \cap V$ and this set is an $l$-dimensional submanifold of $M$. Given $p \in L(C)$, there exists an $l$-dimensional submanifold $L \subset M$ such that $p \in L \subset C$. We choose $L$ small enough, together with an $\epsilon > 0$, such that the normal exponential map $\exp^{TL^\perp}$ restricted to the set of vectors of length $< \epsilon$ is a diffeomorphism onto some neighborhood $V$ of $p$ in $M$ and furthermore $B_\epsilon \subset TM_q$ is a normal ball for every $q \in V$. Clearly $L \subset L(C) \cap V \subset C \cap V$. Now suppose that $q \in (C \cap V) \setminus L$. Then there exist a point $x \in L$ and a segment $qx$ of length $< \epsilon$ perpendicular to $L$. It follows that the geodesic cone $\{\exp_q(\lambda v) : |v| < \epsilon, \exp_q(v) \in L, \lambda \in [0, 1]\}$ contains a non-empty $(l + 1)$-dimensional submanifold of $M$. Since $q \in C$ and $L \subset C$, this cone lies in $C$, we thus get a contradiction to the definition of $l$. This shows that $L = C \cap V$. Hence $L = L(C) \cap V = C \cap V$, and this is an $l$-dimensional submanifold of $M$.

Now let $c\colon [a, b] \to M$ be a unit speed geodesic with $c(a) \in L(C)$ and $c(b) \in C$, where $a < b$. Note that $c([a, b]) \subset C$. There exists an $\epsilon > 0$ such that $B_\epsilon \subset TM_{c(t)}$ is a normal ball for every $t \in [a, b]$. We claim that if $c(r) \in L(C)$ for some

$r \in [a,b)$, then $c(s) \in L(C)$ for all $s \in (r, \min\{r+\epsilon, b\})$. For some neighborhood $V$ of $p := c(r) \in L(C)$ as above, we know that $c(s) \in L(C)$ as long as $c(s) \in V$, in particular $c'(r)$ is tangent to $L(C)$. Let $H$ be an $(l-1)$-dimensional submanifold of $L(C)$ through $p$ orthogonal to $c'(r)$, and let $q := c(t)$ for some $t \in (r, r+\epsilon)$, $t \le b$. It follows as above that the cone $\{\exp_q(\lambda v) : |v| < \epsilon, \exp_q(v) \in H, \lambda \in [0,1]\} \subset C$ contains an $l$-dimensional submanifold of $M$, which in turn contains the open segment $c((r,t))$. This shows the claim. Since $c(a) \in L(C)$, we conclude that $c([a,b)) \subset L(C)$. □

For $C$ and $L(C)$ as in Proposition 5.20, we now define the *border*

$$\mathrm{bd}(C) := C \setminus L(C).$$

Note that $\mathrm{bd}(C)$ equals the topological boundary $\partial C$ of $C$ in $M$ if and only if $\dim(C) = m = \dim(M)$. It can be shown that for every $p \in \mathrm{bd}(C)$, the set $\exp_p^{-1}(L(C)) = \{v \in TM_p : \exp_p(v) \in L(C)\}$ is a relatively open subset of an $l$-dimensional linear subspace of $TM_p$, which we denote by $TC_p$. Furthermore,

$$\exp^{-1}(L(C)) \subset \{w \in TC_p : \langle v_p, w \rangle > 0\}$$

for some unit vector $v_p \in TC_p$, and $v_p$ is unique for a dense set of points in $\mathrm{bd}(C)$. We refer to [ChE1975] or [ChG1972] for the proofs.

**5.21 Proposition** *Let $M$ be a complete Riemannian manifold of non-negative sectional curvature. Let $C \neq \emptyset$ be a compact totally convex subset of $M$ with* $\mathrm{bd}(C) \neq \emptyset$. *Then the function $f := d(\,\cdot\,, \mathrm{bd}(C)) \colon C \to \mathbb{R}$ is concave, thus for every $r \in \mathbb{R}$ the superlevel set $\{f \ge r\}$ is a compact totally convex subset of $M$.*

*Proof*: Since $L(C)$ is itself totally convex, by an approximation argument it suffices to prove that the restriction of $f$ to $L(C)$ is concave. Let $c \colon [0,1] \to L(C)$ be a minimizing geodesic from $x_0$ to $x_1 \neq x_0$, let $\lambda \in (0,1)$, and put $q := c(\lambda)$. Let $p \in \mathrm{bd}(C)$ be a point such that $d(p,q) = f(q)$. Choose segments $px_0$, $px_1$ and a comparison triple $(\bar p, \bar x_0, \bar x_1)$ in $\mathbb{R}^2$ for $(p, x_0, x_1)$. Let $\bar q := (1-\lambda)\bar x_0 + \lambda \bar x_1$ be the point corresponding to $q$. Theorem 5.15 (Toponogov) shows that

$$\gamma := \measuredangle_p(x_0, x_1) \ge \bar\gamma := \measuredangle_{\bar p}(\bar x_0, \bar x_1)$$

(angle comparison ($\mathrm{A}_0$)) and $f(q) = d(p,q) \ge d(\bar p, \bar q)$ (distance comparison ($\mathrm{D}_0$) for the hinge formed by $c$ and the segment $x_0 p$). Note also that $\gamma < \pi$, for $\gamma = \pi$ would imply $p \in L(C)$ by the total convexity of $L(C)$. Furthermore, in case $\bar\gamma = 0$, one of the points $x_0, x_1$ lies on a segment from $p$ to the other and obviously

$$f(q) = d(p,q) = (1-\lambda)d(p,x_0) + \lambda d(p,x_1) \ge (1-\lambda)f(x_0) + \lambda f(x_1).$$

Suppose now that $\bar\gamma > 0$. For $i = 0,1$, let $\bar y_i(\varphi)$ be the point on the half-line $\{\bar x_i + s(\bar p - \bar q) : s \ge 0\}$ with $\measuredangle_{\bar p}(\bar x_i, \bar y_i(\varphi)) = \varphi \ge 0$. Let $u_i \in TC_p$ be the initial vector of $px_i$ with $|u_i| = d(p, x_i)$. Since $\measuredangle(u_0, u_1) = \gamma \in (0, \pi)$, the vectors $u_0$ and $u_1$ span a plane $P$ in $TC_p$. Let $P_i \subset P$ be the closed half-plane bounded by $\mathbb{R}u_i$ that does not contain $u_{1-i}$, and denote by $v_i(\varphi) \in P_i$ the point with $|v_i(\varphi)| = |\bar{y}_i(\varphi) - \bar{p}|$ and $\measuredangle(u_i, v_i(\varphi)) = \varphi$. Since $\exp_p(v_i(0)) = \exp_p(u_i) = x_i \in L(C)$ and $\exp_p^{-1}(L(C))$ is open in $TC_p$, there is a smallest angle $\omega_i > 0$ such that $z_i := \exp_p(v_i(\omega_i)) \notin L(C)$ and thus $z_i \in \mathrm{bd}(C)$. Then $\omega_0 + \gamma + \omega_1 \le \pi$, for otherwise there would exist $\varphi_0, \varphi_1$ such that $\varphi_0 + \gamma + \varphi_1 = \pi$ and $\exp_p(v_i(\varphi_i)) \in L(C)$, which would imply $p \in L(C)$ as above. Let $\bar{z}_i := \bar{y}_i(\omega_i)$. As $\omega_0 + \bar{\gamma} + \omega_1 \le \pi$, it follows readily that

$$f(q) = d(p, q) \ge d(\bar{p}, \bar{q}) \ge (1 - \lambda)\, d(\bar{x}_0, \bar{z}_0) + \lambda\, d(\bar{x}_1, \bar{z}_1).$$

By the generalized hinge version of Toponogov's Theorem (note that we do not know whether $[0, 1] \ni t \mapsto \exp_p(tv_i(\omega_i))$ is minimizing) we get that $d(\bar{x}_i, \bar{z}_i) \ge d(x_i, z_i) \ge f(x_i)$. This gives the result. $\square$

**5.22 Theorem (Gromoll–Meyer 1969, Cheeger–Gromoll 1972)** *Suppose that $M$ is a complete, non-compact $m$-dimensional Riemannian manifold of non-negative sectional curvature. Then $M$ contains a compact totally convex submanifold $\Sigma$ of dimension $l \in \{0, \dots, m-1\}$ such that $M$ is diffeomorphic to the total space of the normal bundle $T\Sigma^\perp$. If the sectional curvature of $M$ is positive, then $M$ is diffeomorphic to $\mathbb{R}^m$.*

The manifold $\Sigma$ is called a *soul* of $M$. The construction of $\Sigma$ results from Proposition 5.19 and Proposition 5.21 and will depend (only) on the choice of the point $p \in M$ in the first of these results. For example, if $M$ is the flat cylinder $S^1 \times \mathbb{R} \subset \mathbb{R}^2 \times \mathbb{R}$, then each of the circles $S^1 \times \{r\}$ is a soul. For $M = \mathbb{R}^m$, every singleton is a soul.

*Proof*: Part I. Proposition 5.19 gives a descending family $(C_r)_{r \in (-\infty, t]}$ of non-empty compact totally convex subsets of $M$, where $C_s$ is the inner parallel set of $C_r$ at distance $\ge s - r$ from $\partial C_r$ ($r < s \le t$), and $C_t$ has empty interior, thus $\dim(C_t) < m$. Put $t_1 := t$ and $l_1 := \dim(C_{t_1})$. If $\mathrm{bd}(C_{t_1}) = \emptyset$, then $\Sigma := L(C_{t_1}) = C_{t_1}$ is already the desired compact submanifold.

If $\mathrm{bd}(C_{t_1}) \neq \emptyset$, then Proposition 5.21 gives a further family of non-empty compact totally convex sets

$$C_r := \{q \in C_{t_1} : d(q, \mathrm{bd}(C_{t_1})) \ge r - t_1\}, \quad r \in (t_1, t_2],$$

where $t_2$ is the minimal $r$ for which $C_r$ has empty interior relative to $L(C_{t_1})$. Then $\dim(C_r) = l_1$ for $r \in [t_1, t_2)$, whereas $l_2 := \dim(C_{t_2}) < l_1$. If $\mathrm{bd}(C_{t_2}) = \emptyset$, then we put $\Sigma := L(C_{t_2}) = C_{t_2}$.

If $\mathrm{bd}(C_{t_2}) \neq \emptyset$, then by iterating the last step a finite number of times we finally arrive at a family $(C_r)_{r \in (-\infty, t_k]}$ of non-empty compact totally convex sets such that, for some $t_1 < \ldots < t_k$ and $m > l_1 > \ldots > l_k \ge 0$, we have that $\dim(C_r) = l_i$

for $r \in [t_i, t_{i+1})$, $\dim(C_{t_k}) = l_k$, $\mathrm{bd}(C_r) \neq \emptyset$ for $r < t_k$, and $\mathrm{bd}(C_{t_k}) = \emptyset$. Then $\Sigma := L(C_{t_k}) = C_{t_k}$ is the desired submanifold.

Part II. We now show that $M$ is diffeomorphic to $T\Sigma^\perp$. Since $\Sigma$ is compact, there exists an $\epsilon > 0$ such that the restriction of the normal exponential map $\exp^\perp$ of $\Sigma$ to $\{u \in T\Sigma^\perp : |u| < 3\epsilon\}$ is a diffeomorphism onto the open $3\epsilon$-neighborhood $U_{3\epsilon}$ of $\Sigma$ in $M$. We will show that $M$ is diffeomorphic to $U_{2\epsilon}$; this suffices since $\{u \in T\Sigma^\perp : |u| < 2\epsilon\}$ is diffeomorphic to $T\Sigma^\perp$.

Let $p \in M \setminus \Sigma$. Then there exists an $r$ such that $p \in \mathrm{bd}(C_r)$. Furthermore, as indicated earlier, there exists a unit vector $v_p \in T(C_r)_p \subset TM_p$ such that $\langle v_p, c'(0)\rangle > 0$ for every geodesic $c\colon [0,1] \to M$ from $p$ to $\Sigma$ of length $d(p,\Sigma)$. We say that $v_p$ *points towards* $\Sigma$. It is not difficult to see that $v_p$ can be extended to a smooth vector field $V_p$ in a neighborhood of $p$ consisting of vectors pointing towards $\Sigma$. Let $V_\Sigma$ denote the gradient field of $3\epsilon - d(\,\cdot\,,\Sigma)$ on $U_{3\epsilon} \setminus \Sigma$. By means of partition of unity we can combine these local fields to a smooth vector field $X$ on $M \setminus \Sigma$ that agrees with $V_\Sigma$ on $U_{2\epsilon} \setminus \Sigma$. Note that every $X_p$ still points towards $\Sigma$. The maximal integral curve $\sigma_p\colon (\alpha_p, \omega_p) \to M \setminus \Sigma$ of $X$ with $\sigma_p(0) = p$ therefore satisfies $d(\sigma(t),\Sigma) < d(\sigma(s),\Sigma)$ for $t > s$. If $p \in U_{2\epsilon}$, then clearly $\omega_p = d(p,\Sigma)$. On the other hand, if $p \in M \setminus U_{2\epsilon}$, then a compactness argument shows that $\sigma_p$ eventually reaches $U_{2\epsilon}$, so that again $\omega_p < \infty$.

Let now $\psi\colon [0,\infty) \to [0,2\epsilon)$ be a smooth function with $\psi' > 0$ such that $\psi(t) = t$ for $t \in [0,\epsilon]$ and $\lim_{t\to\infty}\psi(t) = 2\epsilon$. Define $F\colon M \to U_{2\epsilon}$ such that $F(p) = p$ for $p \in \Sigma$ and $F(p) = \sigma_p(\omega_p - \psi(\omega_p))$ for $p \in M \setminus \Sigma$. Note that $F(p) = p$ for $p \in U_\epsilon \setminus M$. Then $F$ is the required diffeomorphism.

Part III. Suppose now that $M$ has positive sectional curvature. We show that then $\Sigma$ is a singleton $\{z\}$; then $M$ is diffeomorphic to $TM_z = \mathbb{R}^m$.

Suppose to the contrary that there are two distinct points $x_0, x_1 \in \Sigma = L(C_{t_k}) = C_{t_k}$. Let $c\colon [0,1] \to \Sigma$ be a minimizing geodesic from $x_0$ to $x_1$ as in the proof of Proposition 5.21. Recall that $\Sigma$ is of the form $\{q \in C_{t_{k-1}} : d(q, \mathrm{bd}(C_{t_{k-1}})) \geq t_k - t_{k-1}\}$; in the case $k = 1$, this holds for an arbitrary $t_0 < t_1$. Furthermore $\Sigma$ has empty interior relative to $L(C_{t_{k-1}})$. It follows that $d(c(\lambda), \mathrm{bd}(C_{t_{k-1}})) = t_k - t_{k-1}$ for all $\lambda \in [0,1]$. This yields $\gamma = \bar{\gamma}$ in the proof of Proposition 5.21, and one can then conclude that $\sec(P) = 0$ for the plane $P \in TC_p$ (exercise). □

Sharafutdinov [Sh1977] showed that there exists a 1-Lipschitz retraction from $M$ onto $\Sigma$. Using this retraction and Theorem 3.24, Perelman [Pe1994] showed that $\Sigma$ is a singleton already when $\sec(P) > 0$ for all planes $P \subset TM_p$ at *some* point $p \in M$, thus resolving the *Cheeger–Gromoll soul conjecture*.

# Chapter 6

# Volume comparison and applications

## Volume comparison theorems

For $r > 0$ we denote by $V_{m,\kappa}(r)$ the volume of a ball of radius $r$ in $\mathbb{M}^m_\kappa$. Explicitly,

$$V_{m,\kappa}(r) = \omega_{m-1} \int_0^{\inf\{r,D_\kappa\}} \mathrm{sn}_\kappa^{m-1}(t)\,dt,$$

where $\omega_{m-1}$ is the $(m-1)$-dimensional volume of the unit sphere in $\mathbb{R}^m$.

**6.1 Theorem (Günther 1960)** *Let $M$ be a Riemannian $m$-manifold with* $\sec \le \kappa$ *for some $\kappa \in \mathbb{R}$. Let $p \in M$, and let $r > 0$ be such that $B_r \subset TM_p$ is a normal ball. Then*

$$\mathrm{Vol}(B(p,r)) \ge V_{m,\kappa}(r).$$

*Proof*: This follows easily from Corollary 3.19 (exercise). See [Gü1960]. □

For a map $F\colon M \to M'$ between two Riemannian $m$-manifolds and a point $q \in M$ where $F$ is differentiable, $JF(q)$ will denote the *Jacobian* (volume distortion factor) of $F$ at $q$. Thus, for any basis $(b_1,\dots,b_m)$ of $TM_q$,

$$JF(q) = \frac{\sqrt{\det\bigl(\langle dF_q(b_i), dF_q(b_j)\rangle\bigr)}}{\sqrt{\det\bigl(\langle b_i, b_j\rangle\bigr)}}.$$

We define the positive function $J_{m,\kappa}\colon [0,D_\kappa) \to \mathbb{R}$ by

$$J_{m,\kappa}(r) := \frac{\mathrm{sn}_\kappa^{m-1}(r)}{r^{m-1}}$$

for $r \in (0,D_\kappa)$, and $J_{m,\kappa}(0) := 1$. This is the Jacobian of the exponential map $\exp_p\colon T(\mathbb{M}^m_\kappa)_p \to \mathbb{M}^m_\kappa$, for any $p \in \mathbb{M}^m_\kappa$, at any point $v \in T(\mathbb{M}^m_\kappa)_p$ with $|v| = r$.

The following result, as well as Theorem 6.6 below, appeared in [BisC1964].

**6.2 Proposition (Bishop 1964)** *Let $M$ be a Riemannian m-manifold. Suppose that $0 < l \le \infty$ and*

$$\varrho\colon [0,l) \to M, \quad \varrho(r) = \exp_p(ru),$$

*is a unit speed geodesic such that no $\varrho(r)$ with $r \in (0,l)$ is conjugate to $p$ along $\varrho|_{[0,r]}$. Suppose further that* $\mathrm{Ric}(\varrho'(r),\varrho'(r)) \ge \kappa(m-1)$ *for all $r \in [0,l)$, for some constant $\kappa \in \mathbb{R}$. Then $l \le D_\kappa$, and the function*

$$f\colon [0,l) \to \mathbb{R}, \quad f(r) = \frac{J\exp_p(ru)}{J_{m,\kappa}(r)},$$

*satisfies $1 = f(0) \ge f(r) \ge f(s) > 0$ whenever $0 \le r \le s < l$. Moreover, if $f(r) = 1$ for some $r > 0$, then* $\sec(P) = \kappa$ *for every plane $P$ tangent to $\varrho|_{[0,r]}$.*

*Proof*: Clearly $f(0) = 1$, because $d(\exp_p)_0$ is the identity on $TM_p$.

Let now $r \in (0,l)$ be fixed. Assume first that $r$ is also less than $D_\kappa$. We want to show that $f'(r) \le 0$. Let $Y_1,\dots,Y_{m-1}$ be Jacobi fields along $\varrho$ such that $Y_i(0) = 0$ and $(Y_1(r),\dots,Y_{m-1}(r))$ is an orthonormal basis of the normal space $\varrho'(r)^\perp \subset TM_{\varrho(r)}$ (recall Remark 3.10). They are of the form

$$Y_i(t) = d(\exp_p)_{tu}(tv_i)$$

for a basis $(v_1,\dots,v_{m-1})$ of $u^\perp \subset TM_p$ (see the proof of Lemma 3.11). Note that $J\exp_p(tu) = e(t)/\bar e(t)$ for $t \in (0,l)$, where

$$e(t) := \sqrt{\det\bigl(\langle Y_i(t),Y_j(t)\rangle\bigr)} \quad\text{and}\quad \bar e(t) := \sqrt{\det\bigl(\langle tv_i,tv_j\rangle\bigr)} = t^{m-1}\bar e(1)$$

are the $(m-1)$-dimensional volumes of the parallelepipeds spanned by the vectors $Y_i(t)$ and $tv_i$, respectively. Since $Y_1(r),\dots,Y_{m-1}(r)$ are orthonormal, it follows that

$$e'(r) = \frac{1}{2}\det\bigl(\langle Y_i(r),Y_j(r)\rangle\bigr)' = \sum_{i=1}^{m-1}\langle Y_i(r),Y_i'(r)\rangle.$$

Since $Y_i$ is a Jacobi field with $Y_i(0) = 0$, the term $\langle Y_i(r),Y_i'(r)\rangle$ agrees with $I^r(Y_i,Y_i)$, the index form of $\varrho|_{[0,r]}$. Let now $E_i$ denote the parallel vector field along $\varrho$ with $E_i(r) = Y_i(r)$, and put $\lambda := \mathrm{sn}_\kappa(r)$ and $h := \frac{1}{\lambda}\mathrm{sn}_\kappa$. Proposition 3.16 (first index lemma) then yields the inequality

$$\begin{aligned}\langle Y_i(r),Y_i'(r)\rangle &= I^r(Y_i,Y_i)\\ &\le I^r(hE_i,hE_i) = \int^r (h')^2 - h^2R(E_i,\varrho',E_i,\varrho')\,dt.\end{aligned}$$

Using the assumption on the Ricci curvature we deduce that

$$e'(r) \le (m-1)\int_0^r (h')^2 - \kappa h^2\,dt.$$

By the choice of $hE_i$, equality holds for the model space $\mathbb{M}^m_\kappa$, where $e(t)$ becomes $h^{m-1}(t)$. Thus $e'(r) \le (h^{m-1})'(r)$. Since

$$f(t) = \frac{J\exp_p(tu)}{J_{m,\kappa}(t)} = \frac{e(t)}{\bar{e}(t)}\,\frac{t^{m-1}}{\mathrm{sn}_\kappa^{m-1}(t)} = \frac{\lambda^{m-1}}{\bar{e}(1)}\,\frac{e(t)}{h^{m-1}(t)}$$

and $e(r) = 1 = h^{m-1}(r)$, it follows that $f'(r) \le 0$.

To show that $l \le D_\kappa$, suppose to the contrary that $D_\kappa < l$. Then the above argument still shows that $f$ is non-increasing on $[0, D_\kappa)$. However, $J_{m,\kappa}(r) \to 0$ for $r \to D_\kappa$, whereas $J\exp_p(ru) > 0$ for $r \in [0,l)$, because $\varrho$ has no conjugate points. This is a contradiction.

Now suppose that $f(r) = 1$ for some $r > 0$. Then $f$ is constant on $[0,r]$, thus $f'(r) = 0$. It follows that $I^r(Y_i, Y_i) = I^r(hE_i, hE_i)$ in the above argument, for $i = 1, \dots, m-1$. Then Proposition 3.16 shows that $Y_i = hE_i$ on $[0,r]$. By taking linear combinations, we conclude that for every parallel unit normal field $E$ along $\varrho|_{[0,r]}$, the product $\mathrm{sn}_\kappa E$ is a Jacobi field. Thus, by the Jacobi equation,

$$\left\langle \mathrm{sn}_\kappa E, R(\mathrm{sn}_k E, \varrho')\varrho' \right\rangle = \left\langle \mathrm{sn}_\kappa E, -\,\mathrm{sn}''_\kappa E \right\rangle = \kappa\,\mathrm{sn}^2_\kappa\,.$$

It follows that $\sec(\mathrm{span}\{\varrho'(t), E(t)\}) = \kappa$ for all $t \in [0,r]$. □

**6.3 Definition** Suppose that $M$ is complete, and let $p \in M$. For a unit vector $u \in TM_p$, the number

$$t_u := \sup\{t > 0 : d(p, \exp_p(tu)) = t\} \in (0, \infty]$$

is called the *cut value* of $u$. Then

$$\mathrm{Cut}_p := \{t_u u : u \in TM_p,\ |u| = 1,\ t_u < \infty\} \subset TM_p$$

defines the *tangent cut locus* of $p$, and the set $\mathrm{Cut}(p) := \exp_p(\mathrm{Cut}_p)$ of *cut points* of $p$ is the *cut locus* of $p$. The *injectivity radius* of $M$ at $p$ is defined as

$$\mathrm{inj}(p) := d(p, \mathrm{Cut}(p)) = \inf\{t_u : u \in TM_p,\ |u| = 1\},$$

and $\mathrm{inj}_M := \inf\{\mathrm{inj}(p) : p \in M\}$ is the *injectivity radius* of $M$.

**6.4 Lemma** *Suppose that $M$ is complete. Let $c_u : \mathbb{R} \to M$, $c_u(t) := \exp_p(tu)$, be a unit speed geodesic. If the cut value $t_u$ is finite, then* (*at least*) *one of the following holds for $t = t_u$:*

(1) *$c_u(t)$ is conjugate to $p$ along $c_u|_{[0,t]}$,*

(2) *there exists $v \in TM_p$, $|v| = 1$, $v \neq u$, such that $c_v(t) = c_u(t)$.*

*Conversely, if* (1) *or* (2) *holds for some $t \in (0, \infty)$, then $t_u \le t$.*

*Proof*: Exercise. $\square$

**6.5 Proposition** *Suppose that $M$ is complete, and let $p \in M$.*

(1) *The function $u \mapsto t_u$ ($u \in TM_p$, $|u| = 1$) is continuous.*

(2) *The set $U_p := \{tu : 0 \le t < t_u\}$ is open, $\partial U_p = \mathrm{Cut}_p$, and $\exp_p|_{U_p}$ is a diffeomorphism from $U_p$ onto the open set $M \setminus \mathrm{Cut}(p)$.*

(3) $\mathrm{Cut}(p)$ *is a set of measure zero in $M$.*

*Proof*: Exercise. $\square$

**6.6 Theorem (Bishop 1964)** *Let $M$ be a Riemannian m-manifold with* $\mathrm{Ric} \ge (m-1)\kappa g$ *for some constant $\kappa \in \mathbb{R}$. If $p \in M$, and if either $r > 0$ is such that $B_r \subset TM_p$ is a normal ball or $M$ is complete and $r > 0$ is arbitrary, then*

$$\mathrm{Vol}(B(p,r)) \le V_{m,\kappa}(r).$$

*If $M$ is complete and $\kappa > 0$, then* $\mathrm{Vol}(M) \le \mathrm{Vol}(\mathbb{M}^m_\kappa)$*, and equality holds only if $M$ is isometric to $\mathbb{M}^m_\kappa$.*

*Proof*: If $r > 0$ is such that $B_r \subset TM_p$ is a normal ball, then it follows from Proposition 6.2 that $r \le D_\kappa$ and

$$\mathrm{Vol}(B(p,r)) = \int_{B_r} J\exp_p(v)\,dv \le \int_{B_r} J_{m,\kappa}(|v|)\,dv = V_{m,\kappa}(r).$$

Suppose now that $M$ is complete. If $0 < r \le D_\kappa$, then the desired inequality is obtained in the same way, with the only difference that the first integral is taken only over $B_r \cap U_p$; see Proposition 6.5. For $\kappa > 0$ and $r = D_\kappa$, this yields $\mathrm{Vol}(M) \le \mathrm{Vol}(\mathbb{M}^m_\kappa)$, because then $\mathrm{Diam}(M) \le r = \mathrm{Diam}(\mathbb{M}^m_\kappa)$ and $U_p \subset B_r$. In particular, the inequality $\mathrm{Vol}(B(p,r)) \le V_{m,\kappa}(r)$ also holds for $r > D_\kappa$.

Suppose further that $M$ is complete. If $\kappa > 0$ and $\mathrm{Vol}(M) = \mathrm{Vol}(\mathbb{M}^m_\kappa)$, then it follows from the above argument for $r = D_\kappa$ that $U_p = B_r$ and $J\exp_p(v) = J_{m,\kappa}(|v|)$ for all $v \in B_r$ (note that $J_{m,\kappa} > 0$ on $[0,r)$). Now Proposition 6.2 shows that $\sec(P) = \kappa$ for all planes $P \subset TM_p$. Since $p$ was arbitrary, it follows that $M$ is a space form of curvature $\kappa$. As $M$ and $\mathbb{M}^m_\kappa$ have equal volume, they must be isometric (recall Theorem 4.10). $\square$

The following result was observed in [Gro1981a].

**6.7 Theorem (Bishop–Gromov 1981)** *Let $M$ be a complete Riemannian m-manifold with* $\mathrm{Ric} \ge (m-1)\kappa g$ *for some constant $\kappa \in \mathbb{R}$, and let $p \in M$. Then the function*

$$h\colon (0,\infty) \to \mathbb{R}, \quad h(r) = \frac{\mathrm{Vol}(B(p,r))}{V_{m,\kappa}(r)},$$

*satisfies $h(r) \ge h(s)$ whenever $0 < r \le s < \infty$, and $h(r) \to 1$ as $r \to 0$.*

*Proof*: For $r > 0$, let $D(r)$ denote the set of all unit vectors $u \in TM_p$ with $ru \in U_p$, and define

$$a(r) := \int_{D(r)} r^{m-1} J \exp_p(ru)\, du,$$

where the integration is with respect to the canonical measure on the unit sphere in $TM_p$ with total mass $\omega_{m-1}$. Let $b$ be the corresponding function for $\mathbb{M}^m_\kappa$, thus

$$b(r) := \omega_{m-1} r^{m-1} J_{m,\kappa}(r)$$

for $0 < r < D_\kappa$, and $b(r) := 0$ in case $\kappa > 0$ and $r \geq D_\kappa$. Note that

$$\mathrm{Vol}(B(p,r)) = \int_0^r a(t)\, dt \quad \text{and} \quad V_{m,\kappa}(r) = \int_0^r b(t)\, dt.$$

For the monotonicity of $h$, let first $0 < r < s < D_\kappa$. Then

$$\frac{a(s)}{b(s)} = \int_{D(s)} \frac{J \exp_p(su)}{\omega_{m-1} J_{m,\kappa}(s)}\, du \leq \int_{D(s)} \frac{J \exp_p(ru)}{\omega_{m-1} J_{m,\kappa}(r)}\, du.$$

by Proposition 6.2. Since $D(s) \subset D(r)$ (note that $U_p \subset TM_p$ is star-shaped), the last integral is less than or equal to $a(r)/b(r)$. Thus

$$a(r)\, b(s) \geq b(r)\, a(s).$$

Now this inequality holds true also when $s \geq D_\kappa$, for in that case $a(s) = 0$, because $U_p \subset B(0, D_\kappa)$. Hence, for all $s > r > 0$,

$$\int_0^r \int_r^s a(t_1)\, b(t_2)\, dt_2\, dt_1 \geq \int_0^r \int_r^s b(t_1)\, a(t_2)\, dt_2\, dt_1$$

and therefore

$$\int_0^r a \int_0^s b = \int_0^r a \left( \int_0^r b + \int_r^s b \right) \geq \int_0^r b \left( \int_0^r a + \int_r^s a \right) = \int_0^r b \int_0^s a.$$

This gives the result. Clearly $h(r) \to 1$ as $r \to 0+$. $\square$

As a first application of the last two results we give a short proof, due to Shiohama [Shi1983], of Cheng's maximal diameter theorem [Che1975].

**6.8 Theorem (Cheng 1975)** *Let $M$ be a complete Riemannian $m$-manifold with* $\mathrm{Ric} \geq (m-1)\kappa g$ *for some constant $\kappa > 0$, and with (maximal possible) diameter* $\mathrm{Diam}(M) = D_\kappa$. *Then $M$ is isometric to $\mathbb{M}^m_\kappa$.*

*Proof*: For $p \in M$ and $r > 0$, set $h_p(r) := \mathrm{Vol}(B(p,r))/V_{m,\kappa}(r)$. Now fix $p, \bar{p} \in M$ such that $d(p,\bar{p}) = D_\kappa$. Let $r \in (0, D_\kappa)$, and put $\bar{r} := D_\kappa - r$. The balls $B(p,r)$ and $B(\bar{p},\bar{r})$ are disjoint, so

$$\mathrm{Vol}(M) \geq \mathrm{Vol}(B(p,r)) + \mathrm{Vol}(B(\bar{p},\bar{r})) = h_p(r)\, V_{\kappa,m}(r) + h_{\bar{p}}(\bar{r})\, V_{\kappa,m}(\bar{r}).$$

By Theorem 6.7,
$$h_p(r) \geq h_p(D_\kappa) = \frac{\mathrm{Vol}(M)}{\mathrm{Vol}(\mathbb{M}^m_\kappa)}.$$
Likewise, $h_{\bar{p}}(\bar{r}) \geq \mathrm{Vol}(M)/\mathrm{Vol}(\mathbb{M}^m_\kappa)$. Since $V_{\kappa,m}(r)+V_{\kappa,m}(\bar{r}) = \mathrm{Vol}(\mathbb{M}^m_\kappa)$, it follows that equality holds in all of these estimates, in particular $h_p(r) = \mathrm{Vol}(M)/\mathrm{Vol}(\mathbb{M}^m_\kappa)$, independently of $r$. Since $\lim_{r\to 0} h_p(r) = 1$, this shows that $\mathrm{Vol}(M) = \mathrm{Vol}(\mathbb{M}^m_\kappa)$. Hence, by Theorem 6.6, $M$ is isometric to $\mathbb{M}^m_\kappa$. $\square$

## Growth of the fundamental group

We now discuss some results from [Mi1968].

Let $\Gamma = (\Gamma, \cdot)$ be a group with neutral element $e$. Let $A \subset \Gamma$ be a finite set, and suppose that $A$ *generates* $\Gamma$, that is, every element $\gamma \in \Gamma$ can be written as a finite string $\gamma = a_0 \cdots a_n$, where $a_0 = e$ and $a_1, \ldots, a_n \in A \cup \{a^{-1} : a \in A\}$. The smallest number $n \geq 0$ for which such a representation of $\gamma$ exists is called the *word length* of $\gamma$ with respect to $A$ and is denoted by $|\gamma|_A$. Note that $|\gamma|_A = 0$ if and only if $\gamma = e$, and $|\gamma|_A = |\gamma^{-1}|_A$ for all $\gamma \in \Gamma$. If a finite generating set $A \subset \Gamma$ exists, $\Gamma$ is called a *finitely generated group*. Setting
$$d_A(\gamma_1, \gamma_2) := |\gamma_1^{-1}\gamma_2|_A$$
for every pair of elements $\gamma_1, \gamma_2 \in \Gamma$, one obtains a metric on $\Gamma$, the *word metric* with respect to $A$. For $\gamma, \gamma_1, \gamma_2 \in \Gamma$,
$$d_A(\gamma\gamma_1, \gamma\gamma_2) = |(\gamma\gamma_1)^{-1}(\gamma\gamma_2)|_A = |\gamma_1^{-1}\gamma_2|_A = d_A(\gamma_1, \gamma_2),$$
so $d_A$ is left-invariant. If $A'$ is another finite generating set of $\Gamma$ and $L := \max_{a\in A} |a|_{A'}$, then evidently
$$d_{A'}(\gamma_1, \gamma_2) \leq L\, d_A(\gamma_1, \gamma_2) \quad \text{for all } \gamma_1, \gamma_2 \in \Gamma.$$
It follows that any two word metrics on a finitely generated group are bi-Lipschitz equivalent.

**6.9 Definition** The *growth function* of $\Gamma$ with respect to the finite generating set $A$ is defined by
$$w_A(r) := \#\{\gamma \in \Gamma : |\gamma|_A \leq r\}$$
for all integers $r \geq 0$. Then $\Gamma$ has *polynomial growth of degree at most* $k \geq 0$ if there exists a constant $c \geq 1$ such that
$$w_A(r) \leq c r^k \quad \text{for all } r,$$
and $\Gamma$ has *exponential growth* if there exists a constant $b > 1$ such that
$$w_A(r) \geq b^r \quad \text{for all } r.$$

These two properties just depend on $\Gamma$ but not on the generating set: if one of them holds for $A$, then it also holds with respect to any other finite generating set $A'$, only the constant $c$ or $b$ may need to be adjusted. Specifically, if $d_{A'} \le L\, d_A$ as above, then $w_A(r) \le w_{A'}(Lr)$ for all $r$. We remark further that $w_A(r+s) \le w_A(r)w_A(s)$, so the upper exponential bound $w_A(r) \le w_A(1)^r$ is always satisfied.

For example, the free abelian group $\Gamma$ on $k$ generators $a_1, \dots, a_k$ has polynomial growth of degree $k$. For $k \ge 2$, the free group $\Gamma$ on $k$ generators $a_1, \dots, a_k$ has exponential growth.

**6.10 Lemma** *Let $(M, g)$ be a compact connected Riemannian manifold with universal covering $(\tilde{M}, \tilde{g})$, and let $\Gamma \subset \mathrm{Isom}(\tilde{M}, \tilde{g})$ be the group of deck transformations. Choose a compact set $K \subset \tilde{M}$ such that $\bigcup_{\gamma \in \Gamma} \gamma(K) = \tilde{M}$, and let $\delta > 0$. Then the set*

$$A := \{a \in \Gamma : d(K, a(K)) < \delta\}$$

*is finite and generates $\Gamma$; in fact, if $\gamma \in \Gamma$ and $n \ge 1$ is an integer such $d(K, \gamma(K)) < n\delta$, then $|\gamma|_A \le n$.*

In particular, the fundamental group $\pi_1(M) \simeq \Gamma$ is finitely generated.

*Proof*: Since the action of $\Gamma$ on $\tilde{M}$ is properly discontinuous, $A$ is finite. Now let $\gamma \in \Gamma$, and let $n$ be a positive integer such that $d(K, \gamma(K)) < n\delta$. Choose $p_0, \dots, p_n \in \tilde{M}$ such that $p_0 \in K$, $p_n \in \gamma(K)$, and $d(p_{i-1}, p_i) < \delta$ for $i = 1, \dots, n$. By the choice of $K$ there exist $\gamma_0, \dots, \gamma_n \in \Gamma$ such that $\gamma_i(K)$ contains $p_i$, where $\gamma_0 = e$ and $\gamma_n = \gamma$. Setting $a_i := \gamma_{i-1}^{-1}\gamma_i$ for $i = 1, \dots, n$, we get that

$$\gamma = \gamma_1(\gamma_1^{-1}\gamma_2)\cdots(\gamma_{n-1}^{-1}\gamma_n) = a_1 a_2 \cdots a_n.$$

Furthermore, $d(K, a_i(K)) = d(\gamma_{i-1}(K), \gamma_i(K)) \le d(p_{i-1}, p_i) < \delta$, thus $a_i \in A$ and $|\gamma|_A \le n$. □

**6.11 Theorem (Milnor 1968)** *Let $M$ be a compact connected Riemannian manifold with* $\sec < 0$. *Then $\pi_1(M)$ has exponential growth.*

This is no longer true in general if the bound on the sectional curvature is replaced by the assumption $\mathrm{Ric} < 0$: by a result of Lohkamp [Lo1992], [Lo1994], for every $m \ge 3$ there exist constants $a_m > b_m > 0$ such that *every* $m$-dimensional manifold $M$ admits a complete Riemannian metric with $-a_m(m-1)g < \mathrm{Ric} < -b_m(m-1)g$.

*Proof*: We continue with the notation of Lemma 6.10. Let $p \in K$ and $n \ge 1$. If $\gamma \in \Gamma$ and $\gamma(K) \cap B(p, n\delta) \ne \emptyset$, then $d(K, \gamma(K)) < n\delta$ and hence $|\gamma|_A \le n$ by the lemma. Thus $B(p, n\delta)$ is covered by $w_A(n)$ (or less) translates of $K$, all of volume $\mathrm{Vol}(K)$, and so

$$\mathrm{Vol}(B(p, n\delta)) \le w_A(n)\,\mathrm{Vol}(K).$$

Furthermore, since $\sec < 0$ and $M$ is compact, there is a constant $\kappa < 0$ such that $\sec \le \kappa$. By Theorem 4.12 (Hadamard–Cartan), $B_{n\delta} \subset T\tilde{M}_p$ is a normal ball, so $\mathrm{Vol}(B(p, n\delta)) \ge V_{m,\kappa}(n\delta)$ by Theorem 6.1 (Günther). Since $V_{m,\kappa}$ is exponential, $\pi_1(M) \simeq \Gamma$ has exponential growth. □

**6.12 Theorem (Milnor 1968)** *Let $M$ be a complete Riemannian $m$-manifold with* $\mathrm{Ric} \ge 0$. *Then every finitely generated subgroup of $\pi_1(M)$ has polynomial growth of degree at most $m$.*

Milnor conjectured that under the assumptions of the theorem, the fundamental group $\pi_1(M)$ itself is finitely generated.

*Proof*: Let again $(\tilde{M}, \tilde{g})$ denote the universal covering of $(M, g)$, and let $\Gamma \simeq \pi_1(M)$ be the group of deck transformations. Suppose that $\Gamma' \subset \Gamma$ is a non-trivial subgroup with a finite generating set $A \subset \Gamma'$. Fix a reference point $p \in \tilde{M}$. Put $\mu := \max_{a \in A} d(p, a(p))$ and note that

$$d(p, \gamma(p)) \le \mu|\gamma|_A \quad \text{for all } \gamma \in \Gamma'.$$

Let $\epsilon > 0$ be such that $d(p, \gamma(p)) \ge 2\epsilon$ for all $\gamma \in \Gamma' \setminus \{e\}$. For every integer $r \ge 0$, the ball $B(p, \mu r + \epsilon)$ contains all balls $B(\gamma(p), \epsilon)$ with $|\gamma|_A \le r$, and the latter are pairwise disjoint by the choice of $\epsilon$. As there are $w_A(r)$ such $\epsilon$-balls, all of the same volume, we get that

$$w_A(r)\,\mathrm{Vol}(B(p, \epsilon)) \le \mathrm{Vol}(B(p, \mu r + \epsilon)) \le V_{m,0}(\mu r + \epsilon)$$

by Theorem 6.6 (Bishop). It follows that $\Gamma'$ has polynomial growth of degree at most $m$. □

For example, Theorem 6.11 shows again that the torus $M = \mathbb{R}^m/\mathbb{Z}^m$ does not admit a metric with $\sec < 0$ (compare Theorem 4.20).

For another application, consider the Heisenberg group $\tilde{M} = H$ of real upper triangular $3 \times 3$-matrices with ones on the diagonal. The lattice $\Gamma \subset H$ of integer matrices is finitely generated and has polynomial growth of degree (exactly) 4 (exercise). Now Theorem 6.11 and Theorem 6.12 show that the compact 3-manifold $\tilde{M}/\Gamma$ does neither admit a metric with $\sec < 0$ nor a metric with $\mathrm{Ric} \ge 0$.

Regarding Milnor's conjecture, Wilking [Wi2000] proved the following result. Suppose that $\Gamma$ is a group such that for some $m$, every finitely generated subgroup has polynomial growth of degree at most $m$ (like $\Gamma = \pi_1(M)$ in Theorem 6.12). Then $\Gamma$ is finitely generated if and only if every abelian subgroup is finitely generated.

For complete manifolds of non-negative *sectional* curvature, the following holds [Gro1978].

**6.13 Theorem (Gromov 1978)** *For every complete $m$-dimensional Riemannian manifold with* $\sec \ge 0$*, the fundamental group $\pi_1(M)$, as well as every subgroup of $\pi_1(M)$, is generated by $3^m$ (or less) elements.*

*Proof*: As in the previous proofs, let $\tilde{M}$ be the universal Riemannian covering of $M$, and let $\Gamma \simeq \pi_1(M)$ be the group of covering transformations. Suppose that $\Gamma' \subset \Gamma$ is a non-trivial subgroup. Fix a point $p \in \tilde{M}$. The following inductive procedure will produce a finite generating set $\{a_1, \dots, a_k\}$ for $\Gamma'$. First, pick $a_1 \in \Gamma' \setminus \{e\}$ so that $d_1 := d(p, a_1(p)) \le d(p, a(p))$ for all $a \in \Gamma' \setminus \{e\}$, and let $\Gamma_1$ denote the subgroup of $\Gamma'$ generated by $a_1$. If $\Gamma_1 = \Gamma'$, then $k = 1$ and the procedure terminates. Suppose now that elements $a_1, \dots, a_{j-1} \in \Gamma'$ have been chosen and the subgroup $\Gamma_{j-1}$ they generate is still smaller than $\Gamma'$. Then pick $a_j \in \Gamma' \setminus \Gamma_{j-1}$ such that

$$d_j := d(p, a_j(p)) \le d(p, a(p)) \quad \text{for all } a \in \Gamma' \setminus \Gamma_{j-1},$$

and let $\Gamma_j$ denote the subgroup generated by $a_1, \dots, a_j$. For every $i \in \{1, \dots, j-1\}$, we have that $a_i \in \Gamma_i \subset \Gamma_{j-1}$ and $a_j \notin \Gamma_{j-1}$, thus $a_i^{-1} a_j \notin \Gamma_{j-1}$ and

$$d_{ij} := d(a_i(p), a_j(p)) = d(p, (a_i^{-1} a_j)(p)) \ge d_j;$$

furthermore $d_j \ge d_i$ by the choice of $a_i$. It then follows from Theorem 5.15 (Toponogov) that for any choice of segments from $p$ to $a_1(p), \dots, a_j(p)$, the angle $\measuredangle_p(a_i(p), a_j(p))$ is greater than or equal to the corresponding angle in a Euclidean triangle with sides of lengths $d_i \le d_j \le d_{ij}$, which is at least $\frac{\pi}{3}$. Hence, the procedure will terminate after finitely many iterations, for $j = 2, \dots, k$, when $\Gamma_k = \Gamma'$. The number $k$ is no larger than the maximal cardinality of a set of unit vectors in $TM_p$ with mutual angles $\ge \pi/3$, or the maximal number of pairwise disjoint open balls of radius $\frac{1}{2}$ contained in $B_{3/2} \subset TM_p$, which is at most $3^m$. $\square$

## Gromov–Hausdorff convergence

For subsets $A, B$ of a metric space $X = (X, d)$ we denote by

$$N_\delta(A) = \{x \in X : d(x, A) \le \delta\}$$

the closed $\delta$-neighborhood of $A$ and by

$$d_{\mathrm{H}}(A, B) = \inf\{\delta \ge 0 : A \subset N_\delta(B),\ B \subset N_\delta(A)\} \in [0, \infty]$$

the *Hausdorff distance* of $A$ and $B$; $d_{\mathrm{H}}$ defines a metric on the set $\mathscr{C}$ of non-empty, closed and bounded subsets of $X$.

Recall that a metric space $X$ is said to be *precompact* or *totally bounded* if for every $\epsilon > 0$, $X$ can be covered by a finite number of closed balls of radius $\epsilon$. A metric space is compact if and only if it is precompact and complete.

The following result goes back to Blaschke's "Auswahlsatz" (Selection Theorem) for compact convex bodies in $\mathbb{R}^3$, see [Bl1916].

**6.14 Theorem** *Suppose that $X$ is a metric space and $\mathscr{C}$ is the set of non-empty, closed and bounded subsets of $X$, endowed with the Hausdorff metric $d_{\mathrm{H}}$.*

(1) *If $X$ is complete, then $\mathscr{C}$ is complete.*

(2) *If $X$ is compact, then $\mathscr{C}$ is compact.*

*Proof*: For the proof (1), let $(B_i)_{i\in\mathbb{N}}$ be a Cauchy sequence in $\mathscr{C}$. We show that it converges in the Hausdorff distance to the closed set

$$C := \bigcap_{i=1}^{\infty} \overline{\bigcup_{j\geq i} B_j}.$$

Let $\epsilon > 0$. Choose $i_0$ such that $d_{\mathrm{H}}(B_i, B_j) < \epsilon/2$ whenever $i, j \geq i_0$. Suppose that $x \in C$. Since $C \subset \overline{\bigcup_{j\geq i_0} B_j}$, there is an index $j \geq i_0$ such that $d(x, B_j) < \epsilon/2$. Hence

$$d(x, B_i) \leq d(x, B_j) + d_{\mathrm{H}}(B_i, B_j) < \epsilon \quad \text{for all } i \geq i_0.$$

This shows that $C \subset N_\epsilon(B_i)$ for $i \geq i_0$, in particular $C$ is bounded. Now let $x \in B_i$ for some $i \geq i_0$. Pick a sequence $i = i_1 < i_2 < \ldots$ such that $d_{\mathrm{H}}(B_m, B_n) < \epsilon/2^k$ whenever $m, n \geq i_k$, $k \in \mathbb{N}$. Then choose a sequence $(x_k)_{k\in\mathbb{N}}$ such that $x_1 = x$, $x_k \in B_{i_k}$ and $d(x_k, x_{k+1}) < \epsilon/2^k$. As $X$ is complete, the Cauchy sequence $(x_k)$ converges to some point $y$. It follows that

$$d(x, y) = \lim_{k\to\infty} d(x, x_k) \leq \sum_{k=1}^{\infty} d(x_k, x_{k+1}) < \epsilon,$$

and $y$ belongs to $\overline{\bigcup_{l\geq k} B_{i_l}}$ for every $k \in \mathbb{N}$. Thus $y \in C$, in particular $C$ is non-empty, and $d(x, C) < \epsilon$. This shows that $B_i \subset N_\epsilon(C)$ for $i \geq i_0$.

To prove (2), suppose now that $X$ is compact. In view of (1), it suffices to show that $\mathscr{C}$ is precompact. Let $\epsilon > 0$. There exists a finite set $A \subset X$ such that $N_\epsilon(A) = X$. We show that every $B \in \mathscr{C}$ is at Hausdorff distance at most $\epsilon$ of some subset of $A$, namely $A_B := A \cap N_\epsilon(B)$. For every $y \in B$ there exists a point $x \in A$ with $d(x, y) \leq \epsilon$, so $x \in A_B$. This shows that $B \subset N_\epsilon(A_B)$. Since also $A_B \subset N_\epsilon(B)$, we conclude that $d_{\mathrm{H}}(B, A_B) \leq \epsilon$. As there are only finitely many distinct subsets of $A$, it follows that $\mathscr{C}$ is precompact. □

**6.15 Definition** The *Gromov–Hausdorff distance* $d_{\mathrm{GH}}(X, Y) \in [0, \infty]$ of two metric spaces $X, Y$ is defined as the infimum of all $\delta \geq 0$ for which there exist a metric space $(Z, d^Z)$ and subsets $X', Y' \subset Z$ isometric to $X, Y$, respectively, such that $d^Z_{\mathrm{H}}(X', Y') \leq \delta$.

This originates from [Gro1981b].

For example, if both $X$ and $Y$ have diameter less than or equal to $D \geq 0$, then $d_{\mathrm{GH}}(X, Y) \leq \frac{1}{2}D$. For this, let $(Z, d^Z)$ be the union of disjoint isometric copies $X', Y'$ of $X, Y$, where $d^Z(x, y) = d^Z(y, x) = \frac{1}{2}D$ for every pair $(x, y) \in X' \times Y'$.

**6.16 Proposition** (1) *$d_{\mathrm{GH}}$ satisfies the triangle inequality, that is, for all metric spaces $X, Y, Z$,*

$$d_{\mathrm{GH}}(X, Z) \leq d_{\mathrm{GH}}(X, Y) + d_{\mathrm{GH}}(Y, Z).$$

(2) *$d_{\mathrm{GH}}$ defines a metric on the set of isometry classes of compact metric spaces.*

*Proof*: See, for example, Proposition 7.3.16 and Theorem 7.3.30 in [BuBI2001]. □

The following example shows that the second part of Proposition 6.16 does not hold more generally for complete and bounded spaces (as the definition of $(\mathscr{C}, d_{\mathrm{H}})$ might suggest). Let $X$ be the geodesic tree with one central vertex and edges of length $1 - \frac{1}{2}, 1 - \frac{1}{3}, 1 - \frac{1}{4}, \ldots$ attached to it, and define $Y$ similarly, but with an additional segment of length 1. Then $d_{\mathrm{GH}}(X, Y) = 0$ despite $X$ and $Y$ being non-isometric (exercise).

Note also that every compact metric space $X$ admits an isometric embedding into the Banach space $l_\infty$ of bounded sequences with the supremum norm: choose a dense sequence $(x_k)_{k \in \mathbb{N}}$ in $X$ and put $f(x) := (d(x, x_k))_{k \in \mathbb{N}} \in l_\infty$; then

$$\|f(x) - f(y)\|_\infty = \sup_k |d(x, x_k) - d(y, x_k)| \leq d(x, y)$$

due to the triangle inequality, and by taking a sequence $x_{k(1)}, x_{k(2)}, \ldots$ converging to $y$ one sees that equality holds. This holds more generally for all separable metric spaces $X$ (with $f(x) := (d(x, x_k) - d(z, x_k))_{k \in \mathbb{N}}$ for some base point $z \in X$) and is due to Fréchet [Fr1909].

A family $(X_\alpha)_{\alpha \in A}$ of metric spaces, for any index set $A$, is called *uniformly precompact* if $\sup_{\alpha \in A} \mathrm{Diam}(X_\alpha) < \infty$ and for all $\epsilon > 0$ there exists an integer $n(\epsilon)$ such that each $X_\alpha$ can be covered by $n(\epsilon)$ closed balls of radius $\epsilon$.

**6.17 Theorem (Gromov 1981)** *Suppose that $(X_\alpha)_{\alpha \in A}$ is a uniformly precompact family of metric spaces.*

(1) *There exists a compact metric space $Z$ such that each $X_\alpha$ admits an isometric embedding into $Z$.*

(2) *Every sequence in $(X_\alpha)_{\alpha \in A}$ has a subsequence that converges in the Gromov–Hausdorff distance to a compact metric space.*

We follow essentially the original proof from [Gro1981b].

*Proof*: We prove (1). For $i \in \mathbb{N} := \{1, 2, \ldots\}$, put $\epsilon_i := 2^{-i}$ and let $n_i \in \mathbb{N}$ be such that each $X_\alpha$ can be covered by $n_i$ closed balls of radius $\epsilon_i$. Fix a partition $\mathbb{N} = \bigcup_{i=1}^\infty N_i$ and a map $\pi \colon \mathbb{N} \setminus N_1 \to \mathbb{N}$ such that the set $N_i$ has $n_1 n_2 \ldots n_i$ elements and for every $k \in N_i$, $\pi^{-1}\{k\}$ is a subset of $N_{i+1}$ of cardinality $n_{i+1}$. In each $X_\alpha$, construct a sequence $(x_k^\alpha)_{k \in \mathbb{N}}$ in the following way. For $i = 1$, choose $(x_k^\alpha)_{k \in N_1}$ so that the $n_1$ balls $\overline{B}(x_k^\alpha, \epsilon_1)$ cover $X_\alpha$. For $i \geq 1$, if the $n_1 n_2 \dots n_i$ points $x_k^\alpha$ with $k \in N_i$ are chosen, pick $(x_l^\alpha)_{l \in N_{i+1}}$ so that for every $k \in N_i$, the $n_{i+1}$ balls $\overline{B}(x_l^\alpha, \epsilon_{i+1})$ with $l \in \pi^{-1}\{k\}$ cover $\overline{B}(x_k^\alpha, \epsilon_i)$ and are contained in $\overline{B}(x_k^\alpha, 2\epsilon_i)$. This inductive process yields for each $\alpha \in A$ a dense sequence $(x_k^\alpha)_{k \in \mathbb{N}}$ in $X_\alpha$. Let $f_\alpha : X_\alpha \to l_\infty$ be the isometric embedding that maps $x$ to $(d(x, x_k^\alpha))_{k \in \mathbb{N}}$. For $i \in \mathbb{N}$, $k \in N_i$, and $l \in \pi^{-1}\{k\}$,

$$|d(x, x_k^\alpha) - d(x, x_l^\alpha)| \leq d(x_k^\alpha, x_l^\alpha) \leq 2\epsilon_i.$$

Hence, each $f_\alpha(X_\alpha)$ lies in the closed subset $Z \subset l_\infty$ of all sequences $(s_k)_{k \in \mathbb{N}}$ such that $0 \leq s_k \leq \sup_\alpha \operatorname{Diam}(X_\alpha) < \infty$ for all $k \in \mathbb{N}$ and

$$|s_k - s_l| \leq 2\epsilon_i \quad \text{for all } i \in \mathbb{N},\, k \in N_i,\, l \in \pi^{-1}\{k\}.$$

It follows that for all $i \in \mathbb{N}$ and $l \in \bigcup_{j=i+1}^\infty N_j$, there exists a $k \in N_i$ such that $|s_k - s_l| \leq \sum_{j=i}^\infty 2\epsilon_j = 4\epsilon_i$. So the complete subspace $Z \subset l_\infty$ is clearly precompact and thus compact.

(2) follows directly from (1) and the second part of Theorem 6.14. □

**6.18 Theorem (Gromov)** *For $m \geq 2$, $\kappa \in \mathbb{R}$, and $D > 0$, the set $\mathcal{M}$ of isometry classes of compact connected Riemannian m-manifolds $M$ with* $\operatorname{Ric} \geq \kappa(m-1)g$ *and* $\operatorname{Diam}(M) \leq D$ *is uniformly precompact. In particular, every sequence in $\mathcal{M}$ has a subsequence that converges in the Gromov–Hausdorff distance to a compact metric space.*

*Proof*: Let $\epsilon \in (0, D)$, and let $M \in \mathcal{M}$. Suppose that $A \subset M$ is a finite set such that $d(p,q) \geq \epsilon$ for every pair of distinct points $p, q \in A$. Let $z \in A$ be such that $\operatorname{Vol}(B(z, \epsilon/2)) \leq \operatorname{Vol}(B(p, \epsilon/2))$ for all $p \in A$. Then $\operatorname{Vol}(M) \geq |A| \operatorname{Vol}(B(z, \epsilon/2))$, and therefore

$$|A| \leq \frac{\operatorname{Vol}(M)}{\operatorname{Vol}(B(z,\epsilon/2))} = \frac{\operatorname{Vol}(B(z,D))}{\operatorname{Vol}(B(z,\epsilon/2))} \leq \frac{V_{m,\kappa}(D)}{V_{m,\kappa}(\epsilon/2)} =: n(\epsilon)$$

by Theorem 6.7 (Bishop–Gromov). If $A$ is a maximal such set, then the collection $\{B(p,\epsilon) : p \in A\}$ covers $M$ (otherwise there would exist a point $q \in M$ with $d(p,q) \geq \epsilon$ for all $p \in A$, which we could add to $A$). Hence, every $M \in \mathcal{M}$ can be covered by $n(\epsilon)$ open balls of radius $\epsilon$. The second assertion follows from Theorem 6.17. □

**6.19 Remark** Suppose that a sequence of metric spaces $X_i$ converges in the Gromov–Hausdorff distance to a complete metric space $X$. Then the following hold:

(1) If each $X_i$ is a length space, then $X$ is a length space.

(2) If each $X_i$ is proper, then $X$ is proper.

(3) If each $X_i$ is proper and geodesic, then $X$ is proper and geodesic. (This uses (1), (2), and Theorem 1.25).

(4) If each $X_i$ is a proper, geodesic metric space of curvature $\geq \kappa \in \mathbb{R}$ in the sense of Alexandrov, then $X$ has the same properties. (This follows from (3) and Theorem 5.15.)

## Diffeomorphism finiteness

In this last section we will explain how Theorem 6.18 can be used for the proof of a finiteness theorem due to Cheeger [Ch1970]. We start with a result from [Kl1959].

**6.20 Proposition (Klingenberg)** *Let $M$ be a compact connected Riemannian manifold.*

(1) *If $p \in M$, and $q$ is a point in* $\mathrm{Cut}(p)$ *with minimal distance to $p$,*

$$d := d(p,q) = d(p,\mathrm{Cut}(p)) = \mathrm{inj}(p),$$

*then either there exists a minimizing geodesic $\sigma\colon [0,d] \to M$ from $p$ to $q$ along which $q$ is conjugate to $p$, or there exist precisely two distinct minimizing geodesics $\sigma,\tau\colon [0,d] \to M$ from $p$ to $q$, in which case $\sigma'(d) = -\tau'(d)$.*

(2) *If $\lambda > 0$ is such that* $\sec \leq \lambda$, *then either* $\mathrm{inj}_M \geq D_\lambda$, *or there exists a closed geodesic in $M$ of length* $2\,\mathrm{inj}_M$.

*Proof*: Exercise. $\square$

**6.21 Proposition (Cheeger)** *Given $m \geq 2$, $\kappa \in \mathbb{R}$, and $D, V > 0$, there is a constant $l > 0$ such that if $M$ is a compact connected Riemannian $m$-manifold with* $\sec \geq \kappa$, $\mathrm{Diam}(M) \leq D$, *and* $\mathrm{Vol}(M) > V$, *then every non-constant closed geodesic in $M$ has length greater than $l$.*

*Proof*: In the case that $\kappa > 0$ we suppose without loss of generality that $D \leq D_\kappa$. Let $\sigma\colon \mathbb{R} \to M$ be a closed unit speed geodesic with period $b > 0$. Put $p := \sigma(0)$ and $u := \sigma'(0)$. For $r \in (0, D)$ and $\alpha \in (0, \pi/2)$ let

$$W_{r,\alpha} := B(0,D) \setminus \{v : |v| \geq r,\ \measuredangle_p(u,v) \in [0,\alpha] \cup [\pi-\alpha,\pi]\} \subset TM_p.$$

Fix a linear isometry $H\colon TM_p \to T(\mathbb{M}^m_\kappa)_{\bar p}$. Since $V < \mathrm{Vol}(M) \leq V_{m,\kappa}(D)$, we can choose $r, \alpha$ such that $\mathrm{Vol}(\exp_{\bar p}(H(W_{r,\alpha}))) = V$. Now $U_p \not\subset W_{r,\alpha}$ (see Proposition 6.5), for otherwise

$$\mathrm{Vol}(M) = \mathrm{Vol}(\exp_p(W_{r,\alpha})) \leq \mathrm{Vol}(\exp_{\bar p}(H(W_{r,\alpha}))) = V$$

by Corollary 3.19. Since all cut values are less than or equal to $D$, it follows that there exists a minimizing unit speed geodesic $\varrho\colon [0,r] \to M$ with $\varrho(0) = p$ such that $\measuredangle_p(u, \varrho'(0))$ is $\leq \alpha$ or $\geq \pi - \alpha$. By reversing $\sigma$ if necessary, we arrange that $\measuredangle_p(u, \varrho'(0)) \leq \alpha$. Recall that $\alpha < \pi/2$. Now it follows from the generalized hinge version of Toponogov's Theorem that there is a constant $l > 0$, depending only on $\kappa, r, \alpha$ (where $r, \alpha$ depend only on $m, \kappa, D, V$), such that for all $s \in (0, l]$,

$$d(\sigma(s), \varrho(r)) < r = d(p, \varrho(r)).$$

Since $p = \sigma(b)$, this implies that $b > l$. □

**6.22 Theorem (Gromov)** *Given $m \geq 2$ and $\kappa, \varrho > 0$, there exists a constant $\delta > 0$ such that if $M, \bar{M}$ are two compact connected Riemannian m-manifolds with*

$$|\sec_M|, |\sec_{\bar{M}}| \leq \kappa, \quad \mathrm{inj}_M, \mathrm{inj}_{\bar{M}} \geq \varrho, \quad \textit{and} \quad d_{\mathrm{GH}}(M, \bar{M}) < \delta,$$

*then $M$ and $\bar{M}$ are diffeomorphic.*

*Proof*: See Section 8.D in [Gro1999] and the references therein. □

**6.23 Theorem (Cheeger 1970)** *Given $m \geq 2$ and $\kappa, D, V > 0$, there exist only finitely many diffeomorphism classes of compact connected Riemannian m-manifolds $M$ with $|\sec_M| \leq \kappa$, $\mathrm{Diam}(M) \leq D$, and $\mathrm{Vol}(M) \geq V$.*

*Proof*: By Proposition 6.20 and Proposition 6.21 there is a constant $\varrho > 0$ such that $\mathrm{inj}_M \geq \varrho$ for every such $M$. Now the result follows from Theorem 6.18 and Theorem 6.22. □

# Bibliography


[Al1951] A. D. Alexandrov: A theorem on triangles in a metric space and some of its applications (Russian). Trudy Mat. Inst. Steklov 38 (1951), 5–23.

[Al1955] A. D. Alexandrov: Die innere Geometrie der konvexen Flächen. Akademie-Verlag, Berlin 1955.

[Be1955] M. Berger: Sur les groupes d'holonomie homogène des variétés à connexion affine et des variétés riemanniennes. Bull. Soc. Math. France 83 (1955), 279–330.

[Be1962] M. Berger: An extension of Rauch's metric comparison theorem and some applications. Illinois J. Math. 6 (1962), 700–712.

[Bi1911] L. Bieberbach: Über die Bewegungsgruppen der Euklidischen Räume. Math. Ann. 70 (1911), no. 3, 297–336.

[Bi1912] L. Bieberbach: Über die Bewegungsgruppen der Euklidischen Räume. (Zweite Abhandlung.) Die Gruppen mit einem endlichen Fundamentalbereich. Math. Ann. 72 (1912), no. 3, 400–412.

[BisC1964] R. Bishop, R. L. Crittenden: Geometry of manifolds, Academic Press 1964.

[Bl1916] W. Blaschke: Kreis und Kugel. Veit 1916. De Gruyter 1956.

[BrH1999] M. R. Bridson, A. Haefliger: Metric Spaces of Non-Positive Curvature. Grundlehren der Math. Wissenschaften, Vol. 319, Springer 1999.

[BuBI2001] D. Burago, Y. Burago, S. Ivanov: A Course in Metric Geometry. Graduate Studies in Math., Vol. 33, Amer. Math. Soc. 2001.

[BurGP1992] Y. D. Burago, M. Gromov, G. Perelman: A. D. Alexandrov spaces with curvature bounded below. Russian Math. Surveys 47 (1992), no. 2, 1–58.

[Bus1985] P. Buser: A geometric proof of Bieberbach's theorems on crystallographic groups. Enseign. Math. 31 (1985), no. 1-2, 137–145.

[Ca1928] É. Cartan: Leçons sur la géométrie des espaces de Riemann. Gauthier–Villars 1928.

[Ch1970] J. Cheeger: Finiteness theorems for Riemannian manifolds. Amer. J. Math. 92 (1970), no. 1, 61–74.

[ChE1975] J. Cheeger, D. G. Ebin: Comparison Theorems in Riemannian Geometry. North-Holland 1975. AMS Chelsea Publishing 2008.

[ChG1972] J. Cheeger, D. Gromoll: On the structure of complete manifolds of nonnegative curvature. Ann. of Math. 96 (1972), 413–443.

[Che1975] S. Y. Cheng: Eigenvalue comparison theorems and its geometric applications. Math. Z. 143 (1975), 289–297.

[Co1935] S. Cohn-Vossen: Existenz kürzester Wege. Comptes Rendus (Doklady) Acad. Sci. URSS 3 (8) (1935), no. 8 (68), 399–342.

[DeL2016] D. Descombes, U. Lang: Flats in spaces with convex geodesic bicombings. Anal. Geom. Metric Spaces 4 (2016), 68–84.

[DoC1992] M. P. do Carmo: Riemannian Geometry. Birkhäuser 1992, 1993.

[FeN2003] W. Fenchel, J. Nielsen: Discontinuous Groups of Isometries in the Hyperbolic Plane. Edited by Asmus L. Schmidt. De Gruyter Studies in Mathematics 29, De Gruyter 2003.

[Fr1909] M. Fréchet: Les dimensions d'un ensemble abstrait. Math. Ann. 68 (1909-10), no. 2, 145–168.

[GaHL2004] S. Gallot, D. Hulin, J. Lafontaine: Riemannian Geometry. Third Edition. Springer 2004.

[GrKM1975] D. Gromoll, W. Klingenberg, W. Meyer: Riemannsche Geometrie im Großen, 2. Auflage. Lect. Notes in Math. 55, Springer 1975.

[GrM1969] D. Gromoll, W. Meyer: On complete open manifolds of positive curvature. Ann. of Math. 90 (1969), 75–90.

[Gro1978] M. Gromov: Almost flat manifolds. J. Differential Geometry 13 (1978), no. 2, 231–241.

[Gro1981a] M. Gromov: Curvature, diameter and Betti numbers. Comment. Math. Helv. 56 (1981), no. 2, 179–195.

[Gro1981b] M. Gromov: Groups of polynomial growth and expanding maps. Publ. Math. IHES 53 (1981), 53–73.

[Gro1999] M. Gromov: Metric Structures for Riemannian and Non-Riemannian Spaces. With Appendices by M. Katz, P. Pansu, and S. Semmes. Based on the 1981 French original. Birkhäuser 1999.

[Grov1987] K. Grove: Metric Differential Geometry. In: V. L. Hansen (Ed.), Differential Geometry, Proceedings, Lyngby 1985, Lect. Notes in Math. 1263, Springer 1987, pp. 171–227.

[Gü1960] P. Günther: Einige Sätze über das Volumenelement eines Riemannschen Raumes. Publ. Math. Debrecen 7 (1960), 78–93.

[Ho1926] H. Hopf: Zum Clifford–Kleinschen Raumproblem. Math. Ann. 95 (1926), no. 1, 313–339.

[HoR1931] H. Hopf, W. Rinow: Über den Begriff der vollständigen differentialgeometrischen Fläche. Comment. Math. Helv. 3 (1931), 209–225.

[Hu2006] J. H. Hubbard: Teichmüller Theory and Applications to Geometry, Topology, and Dynamics. Vol. 1. Teichmüller Theory. Matrix Editions 2006.

[Jo1995] J. Jost: Riemannian Geometry and Geometric Analysis. Springer 1995.

[Ki1891] W. Killing: Ueber die Clifford-Klein'schen Raumformen. Math. Ann. 39 (1891), no. 2, 257–278.

[Kl1959] W. Klingenberg: Contributions to Riemannian geometry in the large. Ann. of Math. 69 (1959), no. 3, 654–666.

[Ku1955] N. H. Kuiper: On $C^1$-isometric embeddings I, II. Nederl. Akad. Wetensch. Proc. Ser. A. 58 = Indag. Math. 17 (1955), 545–556, 683–689.

[Lo1992] J. Lohkamp: Negatively Ricci curved manifolds. Bull. Amer. Math. Soc. 27 (1992), no. 2, 288–291.

[Lo1994] J. Lohkamp: Metrics of negative Ricci curvature. Ann. of Math. 140 (1994), no. 3, 655–683.

[Mi1968] J. Milnor: A note on curvature and fundamental group. J. Differential Geometry 2 (1968), 1–7.

[Mo1968] G. D. Mostow: Quasi-conformal mappings in $n$-space and the rigidity of hyperbolic space forms. Inst. Hautes Études Sci. Publ. Math. 34 (1968), 53–104.

[Mo1973] G. D. Mostow: Strong rigidity of locally symmetric spaces. Ann. of Math. Studies 78, Princeton Univ. Press 1973.

[My1935] S. B. Myers: Riemannian manifolds in the large. Duke Math. J. 1 (1935), 39–49.

[My1941] S. B. Myers: Riemannian manifolds with positive mean curvature. Duke Math. J. 8 (1941), 401–404.

[Na1954] J. Nash: $C^1$ isometric imbeddings. Ann. of Math. 60 (1954), 383–396.

[Na1956] J. Nash: The imbedding problem for Riemannian manifolds. Ann. of Math. 63 (1956), 20–63.

[ON1966] B. O'Neill: The fundamental equations of a submersion. Michigan Math. J. 13 (1966), no. 4, 459–469.

[ON1983] B. O'Neill: Semi-Riemannian Geometry. With Applications to Relativity. Academic Press 1983.

[Pe1994] G. Perelman: Proof of the soul conjecture of Cheeger and Gromoll. J. Differential Geom. 40 (1994), no. 1, 209–212.

[Pl1991] C. Plaut: Almost Riemannian spaces. J. Differential Geom. 34 (1991), no. 2, 515–537.

[Pl1996] C. Plaut: Spaces of Wald–Berestovskii curvature bounded below. J. Geom. Anal. 6 (1996), no. 1, 113–134.

[Pr1942] A. Preissmann: Quelques propriétés globales des espaces de Riemann. Comment. Math. Helv. 15 (1942–1943), 175–216.

[Ra1951] H. E. Rauch: A contribution to differential geometry in the large. Ann. of Math. 54 (1951), 38–55.

[Sh1974] V. A. Sharafutdinov: Complete open manifolds of nonnegative curvature. Siberian Math. J. 15 (1974), no. 1, 126–136.

[Sh1977] V. A. Sharafutdinov: The Pogorelov–Klingenberg theorem for manifolds homeomorphic to $R^n$. Siberian Math. J. 18 (1977), no. 4, 649–657.

[Shi1983] K. Shiohama: A sphere theorem for manifolds of positive Ricci curvature. Trans. Amer. Math. Soc. 275 (1983), no. 2, 811–819.

[Sy1936] J. L. Synge: On the connectivity of spaces of positive curvature. Quart. J. Math. (Oxford Series) 7 (1936), 316–320.

[Te1940] O. Teichmüller: Extremale quasikonforme Abbildungen und quadratische Differentiale. Abh. Preuß. Akad. Wiss. Math.-Naturw. Kl. 22 (1940), 1–197.

[Th1997] W. P. Thurston: Three-Dimensional Geometry and Topology. Vol. 1. Edited by Silvio Levy. Princeton Univ. Press 1997.

[Ti1908] H. Tietze: Über die topologischen Invarianten mehrdimensionaler Mannigfaltigkeiten. Monatsh. Math. Phys. 19 (1908), no. 1, 1–118.

[To1957] V. A. Toponogov: On convexity of Riemannian spaces of positive curvature (Russian). Dokl. Akad. Nauk SSSR 115 (1957), 674–676.

[To1958] V. A. Toponogov: Riemannian spaces having their curvature bounded below by a positive number (Russian). Dokl. Akad. Nauk SSSR 120 (1958), 719–721.

[To1959] V. A. Toponogov: Riemann spaces with curvature bounded below (Russian). Uspehi Mat. Nauk 14 (1959), no. 1, 87–130.

[We1968] A. Weinstein: A fixed point theorem for positively curved manifolds. J. Math. Mech. 18 (1968/1969) 149–153.

[Wi2000] B. Wilking: On fundamental groups of manifolds of nonnegative curvature. Differential Geom. Appl. 13 (2000), no. 2, 129–165.

[Wo2011] J. A. Wolf: Spaces of Constant Curvature. Sixth Edition. AMS Chelsea Publishing 2011.