

Riemannian and Metric Geometry

Urs Lang
ETH Zürich

Preliminary and incomplete version
January 5, 2026

Contents

1	Riemannian manifolds	1
1.1	Riemannian metrics and distance functions	1
1.2	The Levi-Civita connection	3
1.3	Vector fields along maps	7
1.4	Geodesics and the exponential map	10
1.5	The Hopf–Rinow theorem	14
1.6	Geodesic metric spaces	16
2	Curvature	19
2.1	The curvature tensor	19
2.2	Sectional curvature	22
2.3	Ricci and scalar curvature	23
2.4	Curvature of submanifolds	29
2.5	Riemannian products	32
3	Jacobi fields	35
3.1	Second variation of arc length	35
3.2	Jacobi fields	37
3.3	Conjugate points	39
3.4	The Rauch comparison theorem	42
3.5	Focal points and the Rauch–Berger theorem	44
4	Riemannian submersions and coverings	49
4.1	Riemannian submersions	49
4.2	Curvature of Riemannian submersions	51
4.3	Riemannian coverings and space forms	54
4.4	Hadamard manifolds	58
4.5	Isometries of Hadamard manifolds	60
5	Triangle comparison	63
5.1	Some model space geometry	63
5.2	Alexandrov comparisons	65

5.3	Toponogov's theorem	68
5.4	CAT(0) spaces	71
6	Volume comparison and applications	77
6.1	Volume comparison theorems	77
6.2	Growth of the fundamental group	82
6.3	Gromov–Hausdorff convergence	86
6.4	Diffeomorphism finiteness	90
7	Open manifolds of non-negative curvature	93
7.1	Busermann functions	93
7.2	The soul construction	95
7.3	The Sharafutdinov retraction	101
7.4	Proof of the Cheeger–Gromoll conjecture	103
8	Minimal entropy and Mostow's rigidity theorem	107
8.1	The boundary at infinity	107
8.2	Visual metrics	111
8.3	Quasi-isometries	115
8.4	Entropy and conformal densities	118
8.5	The barycenter method	122
8.6	Rigidity results	127
	Bibliography	135

Chapter 1

Riemannian manifolds

1.1 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^∞ . The tangent and cotangent bundles of M are denoted by TM and TM^* , respectively, the vector spaces of C^∞ sections by $\Gamma(TM)$ (vector fields) and $\Gamma(TM^*)$ (1-forms). For integers $r, s \geq 0$,

$$T_{r,s}M = \underbrace{TM \otimes \dots \otimes TM}_r \otimes \underbrace{TM^* \otimes \dots \otimes TM^*}_s$$

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

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

will be identified with the vector space of multilinear functions

$$\underbrace{TM_p^* \times \dots \times TM_p^*}_r \times \underbrace{TM_p \times \dots \times TM_p}_s \rightarrow \mathbb{R}.$$

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

$$T: (\Gamma(TM^*))^r \times (\Gamma(TM))^s \rightarrow C^\infty(M)$$

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

Definition 1.1 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 $p \in M$,

$$g_p: TM_p \times TM_p \rightarrow \mathbb{R}$$

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

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

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

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

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

Remark 1.2 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 $F: M \rightarrow \bar{M}$ 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 .

Definition 1.3 Two Riemannian manifolds (M, g) and $(\bar{M}, \bar{g})$ are called *isometric* if there exists a diffeomorphism $F: M \rightarrow \bar{M}$ such that $F^*\bar{g} = g$; then F is called an *isometry* from (M, g) to $(\bar{M}, \bar{g})$. A smooth map $F: M \rightarrow \bar{M}$ 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^∞ Riemannian manifold (M, g) of dimension m admits an isometric C^∞ embedding into the Euclidean space $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\text{eucl}})$ 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.

Theorem 1.4 (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 $c: [a, b] \rightarrow M$ 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 joined by a piecewise C^1 curve $c: [a, b] \rightarrow M$, 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 (ϕ, U) be a chart of M with $p \in U$ and $\phi(p) = 0$. Let $\epsilon > 0$ be such that the ball $\bar{B}_\epsilon := \{x \in \mathbb{R}^m : |x|_{\text{euc}} \leq \epsilon\}$ is contained in $\phi(U)$, and consider the compact set $K := \phi^{-1}(\bar{B}_\epsilon)$. Since M is Hausdorff, K is closed in M , and we can fix $\epsilon > 0$ so that $q \notin \phi^{-1}(B_\epsilon)$, where B_ϵ is the open ball. Let now $c: [a, b] \rightarrow M$ 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 $c(s) \notin \phi^{-1}(B_\epsilon)$. Hence, $\bar{c} := \phi \circ c|_{[a, s]}$ is a piecewise C^1 curve in $\bar{B}_\epsilon$ connecting 0 to a boundary point of $\bar{B}_\epsilon$. Put $\bar{g} := (\phi^{-1})^*g$ on $\phi(U)$. Since $\bar{B}_\epsilon$ is compact, there exists a constant $\lambda > 0$ such that $|\bar{v}|_{\bar{g}_x} \geq \lambda|v|_{\text{euc}}$ for all $v \in \mathbb{R}^m$ and $x \in \bar{B}_\epsilon$. Now

$$L(c) \geq L(c|_{[a, s]}) = L_{\bar{g}}(\bar{c}) \geq \lambda L_{\text{euc}}(\bar{c}) \geq \lambda\epsilon,$$

independently of the choice of c , thus $d(p, q) \geq \lambda\epsilon > 0$.

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

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

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 $\bar{g} = (\phi^{-1})^*g$, thus

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

These inclusions show that the two topologies agree. $\square$

1.2 The Levi-Civita connection

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

- (1) $\nabla_f X Y = f \nabla_X Y$,
- (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 ∇ 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 $\nabla \cdot Y: \Gamma(TM) \rightarrow \Gamma(TM)$ 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 ∇ , 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 ∇ 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, $A_i = \frac{\partial}{\partial \phi^i}$ for a chart (ϕ, U) .) The *Christoffel symbols* of ∇ 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 $i, j, k \in \{1, \dots, m\}$.

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

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

Proof: By property (2) in Definition 1.5,

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

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. $\square$

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, $A_i = \frac{\partial}{\partial \phi^i}$ and $\alpha^j = \partial \phi^j$.) 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 ∇ with respect to the given frame field $(A_1, \dots, A_m)$. For X and Y as above, $\alpha^i(X) = X^i$, thus

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

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

Remark 1.7 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 $c: (-\epsilon, \epsilon) \rightarrow M$ with $c(0) = p$ and $c'(0) = X_p$, because $X_p(Y^k) = dY_p^k(X_p) = (Y^k \circ c)'(0)$.

Definition 1.8 Let M be a smooth manifold with a connection ∇ on TM .

- (1) The map $T: \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(TM)$,

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

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

- (2) Given a Riemannian metric $g = \langle \cdot, \cdot \rangle$ on M , the connection ∇ 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 $f \in C^\infty(M)$,

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

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 .

Theorem 1.9 (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 ∇ , is called the *Levi-Civita connection* of $g = \langle \cdot, \cdot \rangle$ and is characterized by the property that the *Koszul formula*

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

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

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

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 $D: \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(TM)$ 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,

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

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

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

Solving this equation for Γ_{ij}^k , we obtain the expression

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

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

1.3 Vector fields along maps

Let $F: N \rightarrow M$ be a smooth map. By a *vector field Y along F* we mean a smooth map $Y: N \rightarrow TM$ 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

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

of a curve $c: I \rightarrow M$ with parameter t .

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

- (1) $\nabla_{fX}^F Y = f \nabla_X^F Y$,
- (2) $\nabla_X^F(fY) = X(f)Y + f \nabla_X^F Y$,
- (3) $(\nabla_X^F(Z \circ F))_p = \nabla_{F_*X_p} Z$,

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 ∇^F is called the *connection along F induced by ∇* . 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 Γ_{ij}^k denote the respective Christoffel symbols of ∇ . Let $V \subset N$ be an open set such that $F(V) \subset U$.

Lemma 1.11 *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*

$$\nabla_X^F Y = \sum_{k=1}^m \left(X(Y^k) + \sum_{i,j=1}^m X^i Y^j (\Gamma_{ij}^k \circ F) \right) (A_k \circ F)$$

on V .

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

$$\nabla_X^F Y = \sum_{j=1}^m \nabla_X^F(Y^j(A_j \circ F)) = \sum_{j=1}^m (X(Y^j)(A_j \circ F) + Y^j \nabla_X^F(A_j \circ F))$$

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

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

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

Proposition 1.12 *Let $F: N \rightarrow M$, ∇ , and ∇^F be given as above.*

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

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

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

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

$$X\langle V, W \rangle = \langle \nabla_X^F V, W \rangle + \langle V, \nabla_X^F W \rangle$$

*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 ∇^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

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

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 $\nabla_Y^F F_* X$ is computed similarly. Furthermore, one can check that

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

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

$$(V, W) \mapsto X\langle V, W \rangle - \langle \nabla_X^F V, W \rangle - \langle V, \nabla_X^F W \rangle$$

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 ∇ and the relation $\nabla_{F_*X} A_k = \nabla_X^F (A_k \circ F)$. $\square$

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 $X = \frac{\partial}{\partial \phi^i}$ or $X = \frac{d}{dt}$ (in the case that $N = I \subset \mathbb{R}$), we write $D_X Y$ as

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

respectively. A vector field Y along a curve $c: I \rightarrow M$ is called *parallel* if

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

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

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

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

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

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

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

Proof: This follows from the existence and uniqueness of solutions for ordinary differential equations. $\square$

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

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

on $[a, b]$. In general, P^c depends on c and not only on the endpoints of c . The *holonomy group* 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 $c : [a, b] \rightarrow M$ with $c(a) = c(b) = p$. Most often, Hol_p acts transitively on the unit sphere $S(1) \subset TM_p$. This statement is made precise in Berger's holonomy theorem [Ber1955].

1.4 Geodesics and the exponential map

Definition 1.14 A curve $c : I \rightarrow (M, g)$ 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 $c : [a, b] \rightarrow M$ be any smooth curve. Consider a smooth variation $\gamma : (-\epsilon, \epsilon) \times [a, b] \rightarrow M$ 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 $V_s := V(s, \cdot)$.

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

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

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

$$\frac{\partial}{\partial s} |\dot{\gamma}_s(t)| = \frac{1}{2|\dot{\gamma}_s(t)|} \frac{\partial}{\partial s} \langle \dot{\gamma}_s(t), \dot{\gamma}_s(t) \rangle = \frac{1}{|\dot{\gamma}_s(t)|} \left\langle \frac{D}{\partial s} \dot{\gamma}_s(t), \dot{\gamma}_s(t) \right\rangle,$$

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,

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

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

The variation γ is called *proper* if $\gamma_s(a) = c(a)$ and $\gamma_s(b) = c(b)$ for all s . It follows that a C^∞ curve $c : [a, b] \rightarrow M$ 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 γ of c .

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

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

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*.

Proposition 1.16 (1) *For every $v \in TM$ there exists a geodesic $c_v: (\alpha_v, \omega_v) \rightarrow M$ defined on a maximal interval (α_v, ω_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 $W \rightarrow M, (v, t) \mapsto c_v(t)$, is C^∞ .*

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. $\square$

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 .

Definition 1.17 The C^∞ map $\exp: \Omega \rightarrow M$ 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

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

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 $TS^1 = \mathbb{R}i$, and $\exp_1(ti) = \cos(t) + \sin(t)i = e^{ti}$ 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 $c: \mathbb{R} \rightarrow G$ 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 $G \subset GL(n, \mathbb{C})$, the former is given by the matrix exponential function $A \mapsto e^A$.

Remark 1.18 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 $v \in TM_p$,

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

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

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

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

around p .

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

$$g_{ij}(p) = \delta_{ij}, \quad \frac{\partial g_{ij}}{\partial \phi^k}(p) = 0, \quad \text{and} \quad \Gamma_{ij}^k(p) = 0$$

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

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

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

thus $\dot{x}^i(t) = H(v)^i$, and the geodesic equation for $t \mapsto c_v(t)$ reads

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

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

$$\frac{\partial g_{ij}}{\partial \phi^k} = \sum_{l=1}^m (\Gamma_{ki}^l g_{lj} + g_{il} \Gamma_{kj}^l)$$

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

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

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

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 $\gamma_s: [0, 1] \rightarrow M$, $\gamma_s(t) := \exp_p(t(v+sw))$, where $\gamma_0 = c_v|_{[0,1]}$. By the first variation formula, Theorem 1.15,

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

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

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

this gives the result. $\square$

Proposition 1.21 *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 $\gamma: [a, b] \rightarrow M$ 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 γ 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 $\gamma: [a, b] \rightarrow M$ be given as in (1). There exists a maximal $s \in (a, b]$ such that $\gamma([a, s]) \subset \exp_p(\bar{B}_{|v|})$ (compare the proof of Theorem 1.4), and the piecewise C^1 curve $\beta: [a, s] \rightarrow \bar{B}_{|v|}$ defined by $\beta(t) := \exp_p^{-1}(\gamma(t))$ for all $t \in [a, s]$ connects 0 to a boundary point of $\bar{B}_{|v|} \subset TM_p$. 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 $T_{\beta(t)} := d(\exp_p)_{\beta(t)}$,

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

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

We prove (2). Suppose that $\varrho \leq d(p, q) < \infty$. Choose v with $|v| = \varrho$ such that $q' := \exp_p(v)$ minimizes the distance to q . Let $\epsilon > 0$, and let $\gamma: [a, b] \rightarrow M$ 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

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

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

Corollary 1.22 *Every geodesic $c: I \rightarrow M$ 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 $(\pi|_\Omega, \exp): \Omega \rightarrow M \times M$ is regular at $0_p \in TM_p$, because the projection $\pi: TM \rightarrow M$ maps 0_q to q for all q , and by Remark 1.18. It follows that there is an $\epsilon > 0$ such that $\exp_q|_{B_\epsilon}: B_\epsilon \rightarrow \exp_q(B_\epsilon)$ 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. $\square$

1.5 The Hopf–Rinow theorem

The following fundamental result was established in [HoR1931] for surfaces, and it was then observed in [My1935] that the generalization to higher dimensions is immediate.

Theorem 1.23 (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; such manifolds are called *closed*. By contrast, a complete non-compact Riemannian manifold is called *open*.

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 ϱ 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 $c: [0, \infty) \rightarrow M$ emanating from p such that $|\dot{c}| \equiv 1$ and $c(\varrho) = q'$. Put

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

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 $\sigma: [s, s + \varrho'] \rightarrow M$ of length ϱ' 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 σ 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 $\bar{B}(p, r) \subset \exp_p(\bar{B}_r)$ for all $r > 0$, and the reverse inclusion is clear. Hence, since $\bar{B}_r$ is compact and $\exp_p$ is continuous, $\bar{B}(p, r)$ 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 $c_v: [0, \omega_v) \rightarrow M$ be the maximal geodesic ray with initial 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 γ 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

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

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. $\square$

1.6 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 $c: I \rightarrow X$ 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 $t_0 \leq t_1 \leq \dots \leq t_k$ in I . The curve c has *constant speed* or is *parametrized proportionally to arc length* if there exists a constant $\lambda \geq 0$ such that $L(c|_{[t, t']}) = \lambda|t - t'|$ whenever $t, t' \in I$ and $t \leq t'$, 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 $c: [0, 1] \rightarrow X$ from p to q .

Definition 1.24 Let (X, d) be a metric space. A curve $c: I \rightarrow X$ 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 $c: [0, 1] \rightarrow X$ from p to q .

Evidently every geodesic space is also a length space. A partial converse will be established below. We use the following lemma.

Lemma 1.25 Let X be a length space, and suppose that for some $x \in X$ and $R > 0$ the ball $\bar{B}(x, r)$ is compact for every $r \in (0, R)$. Then for every $y \in \bar{B}(x, R)$ there exists a minimizing geodesic $c: [0, d(x, y)] \rightarrow X$ from x to y .

For illustration in a simple case, consider $\mathbb{R}^2 \setminus \{(0, 0)\} \subset \mathbb{R}^2$ with the induced metric and a ball $\bar{B}((0, R), R)$.

Proof: Given $y \in \bar{B}(x, R)$, there is a sequence of unit speed (and hence 1-Lipschitz) curves $c_i: [0, l_i] \rightarrow X$ from x to y such that $l_i \rightarrow l := d(x, y)$ as $n \rightarrow \infty$. For every $r \in (0, l)$, the sequence $(c_i(r))$ lies in the compact ball $\bar{B}(x, r)$. Furthermore, $d(c_i(l), y) \leq l_i - l \rightarrow 0$. Now an Arzelà–Ascoli type (diagonal sequence) argument shows that some subsequence of $(c_i|_{[0, l]})$ converges pointwise (in fact, uniformly) to a 1-Lipschitz curve $c: [0, l] \rightarrow X$ with $c(0) = x$ and $c(l) = y$. This is the desired minimizing geodesic joining x and y . $\square$

We now prove a slight reformulation of the result from [Co1935].

Theorem 1.26 (Cohn-Vossen 1935) *Let X be a locally compact length space. Then the following are equivalent:*

- (1) *for some point $x \in X$, every minimizing geodesic $c: [0, 1) \rightarrow X$ with $c(0) = x$ can be extended to $[0, 1]$;*
- (2) *every curve $c: [0, 1) \rightarrow X$ of finite length can be extended to $[0, 1]$;*
- (3) *X is complete;*
- (4) *X is proper (that is, every closed bounded subset of X is compact).*

If these conditions hold, then X is a geodesic space.

Proof: The implications (4) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (1) hold in every metric space. Furthermore, by Lemma 1.25, every proper length space is geodesic.

Suppose now that X is a locally compact length space such that (1) holds. For this distinguished point $x \in X$, put $I := \{r \geq 0 : \bar{B}(x, r) \text{ is compact}\}$. Note that I is an interval containing 0. Since X is locally compact, I is open relative to $[0, \infty)$: for $r \in I$, $\bar{B}(x, r)$ can be covered by finitely many open balls $B(x_i, \epsilon_i)$ such that the corresponding closed balls $\bar{B}(x_i, \epsilon_i)$ are compact; then their union is compact and contains $\bar{B}(x, r + \delta)$ for some $\delta > 0$.

To prove that I is also closed, suppose that $[0, R) \subset I$ for some $R > 0$, and let (y_n) be a sequence in $\bar{B}(x, R)$. By Lemma 1.25 there is a minimizing geodesic $c_n: [0, 1] \rightarrow X$ from x to y_n . For every $t \in [0, 1)$, the sequence $(c_n(t))$ lies in the compact ball $\bar{B}(x, tR)$. Similarly as in the proof of the lemma, for some subsequence $(c_{n(i)})$, the restrictions $c_{n(i)}|_{[0, 1)}$ converge pointwise (and uniformly on compact sets) to a minimizing geodesic $c: [0, 1) \rightarrow X$. By (1), c extends to $[0, 1]$, and clearly $y_{n(i)} = c_{n(i)}(1) \rightarrow c(1)$ as $i \rightarrow \infty$. This shows that $\bar{B}(x, R)$ is compact, hence $R \in I$.

Consequently, $I = [0, \infty)$, and (4) holds. This completes the proof. $\square$

Theorem 1.23 asserts in particular that every complete and locally compact length space is geodesic. Neither completeness nor local compactness can be dropped:

Example 1.27 The set $\mathbb{R}^2 \setminus \{0\}$, with the metric induced from $\mathbb{R}^2$, is an incomplete, locally compact length space that is not geodesic. The complete length space obtained from a sequence of disjoint copies of $[0, 1 + 1/n]$ ($n \in \mathbb{N}$) by gluing all left endpoints to a point x and all right endpoints to a point y is not locally compact and not geodesic, because there is no segment of length $d(x, y) = 1$ joining x and y .

Regarding the first item in Theorem 1.23, we point out that the continuous extension of a (non-minimizing) geodesic $c: [0, 1) \rightarrow X$ to $[0, 1]$ may fail to be a geodesic, even if X is a compact geodesic space. This stands in contrast to the situation in Riemannian geometry.

Example 1.28 Put $S := \{1 - 2^{-n} : n = 0, 1, 2, \dots\}$. Let $X \subset \mathbb{R}^2$ be the union of $[0, 1] \times \{0\}$ with the graph $\{(t, d(t, S)) : t \in [0, 1]\}$ of the distance function to S , and endow X with the induced length metric. Define $c: [0, 1] \rightarrow X$ by $c(t) := (t, d(t, S))$. Then $c|_{[0, 1]}$ is a geodesic, whereas no subsegment $c|_{[t', 1]}$ of c with $t' < 1$ is minimizing.

Chapter 2

Curvature

2.1 The curvature tensor

Definition 2.1 Let $(M, g = \langle \cdot, \cdot \rangle)$ be a Riemannian manifold with Levi-Civita connection D . The map $R: \Gamma(TM)^3 \rightarrow \Gamma(TM)$ 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 ∇ on TM is defined analogously.

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

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

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

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

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

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

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

$$R(A_k, A_l)A_j = \sum_{i=1}^m R_{jkl}^i A_i,$$

$$R_{jkl}^i := A_k(\Gamma_{lj}^i) - A_l(\Gamma_{kj}^i) + \sum_{s=1}^m (\Gamma_{lj}^s \Gamma_{ks}^i - \Gamma_{kj}^s \Gamma_{ls}^i).$$

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

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

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

Proposition 2.2 (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

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

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.

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

$$R^F(X, Y)W := \nabla_X^F \nabla_Y^F W - \nabla_Y^F \nabla_X^F W - \nabla_{[X, Y]}^F W \in \Gamma(F^*TM)$$

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 $A_1, \dots, A_m$ 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 $\nabla_Y^F(A \circ F) = \nabla_{F_*Y} A = \sum_j Y^j B_j \circ F$, hence

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

and the second sum equals

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

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

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

From these identities, it follows that

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

This gives the result. $\square$

2.2 Sectional curvature

Definition 2.4 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

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

Remark 2.5 For $p \in M$, let $G_2(TM_p)$ denote the set of all linear 2-planes in TM_p . The sectional curvature function $\sec_p: G_2(TM_p) \rightarrow \mathbb{R}$, 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

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

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 $K: M \rightarrow \mathbb{R}$, $K(p) = \sec(TM_p)$, is the *Gauss curvature* of M . In a chart (ϕ, U) ,

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

Definition 2.6 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

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

(exercise). We shall see that the model spaces

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

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 4.13).

2.3 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:

$$T: V^* \times V \rightarrow \mathbb{R} \quad \text{and} \quad \tilde{T}: V \rightarrow V,$$

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

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

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

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

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*

$$C = C_{k,l}: V_{r,s} \rightarrow V_{r-1,s-1}$$

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 $T_{bc}^a = \sum_i T_{bci}^a$.

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

$$C: V_{r,s} \rightarrow V_{r-n,s-n}$$

is a composition of n such (simple) contractions.

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

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

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

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

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

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

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

In a chart (ϕ, U) ,

$$\text{Ric}|_U = \sum_{i,j} R_{ij} d\phi^i \otimes d\phi^j, \quad R_{ij} = \sum_k R_{ikj}^k = R_{ji}.$$

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

$$\text{Ric}'\left(\frac{\partial}{\partial \phi^j}\right) = \sum_i R_j^i \frac{\partial}{\partial \phi^i}, \quad R_j^i = \sum_k g^{ik} R_{kj}.$$

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

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

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

In a chart (ϕ, U) ,

$$\text{scal}|_U = \sum_i R_i^i = \sum_{i,j} g^{ij} R_{ji}.$$

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 2.16), then

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

Next we discuss tensor derivations.

Proposition 2.9 Let ∇ be a connection on TM . There exists a unique family of bilinear operators

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

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

- (1) $\nabla_X^{0,0}f = X(f) = df(X)$,
- (2) $\nabla_X^{1,0}Y = \nabla_X Y$,
- (3) $X(\theta(Y)) = (\nabla_X^{0,1}\theta)(Y) + \theta(\nabla_X Y)$, and
- (4) $\nabla_X^{r+r', s+s'}(T \otimes T') = (\nabla_X^{r,s}T) \otimes T' + T \otimes (\nabla_X^{r',s'}T')$.

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) $C(\nabla_X^{r,s}T) = \nabla_X^{r-n, s-n}C(T)$ for any n -fold contraction $C: \Gamma(T_{r,s}M) \rightarrow \Gamma(T_{r-n, s-n}M)$.

Proof (sketch): Properties (1)–(3) determine $\nabla^{0,0}, \nabla^{1,0}, \nabla^{0,1}$. It is straightforward to check that $(\nabla_X^{0,1}\theta)(fY) = f(\nabla_X^{0,1}\theta)(Y)$ for $f \in C^\infty(M)$, so indeed $\nabla_X^{0,1}\theta \in \Gamma(T_{0,1}M)$. Property (6) follows directly from (1) and (4) since $fT = f \otimes T$. Furthermore, (6) implies that for a fixed X , the operators $\nabla_X^{r,s}$ are locally defined (as for ∇_X). Now one can use a local representation of T and (1)–(4) to show that $\nabla_X^{r,s}T$ 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

$$\nabla_X^{1,1}(Y \otimes \omega) = (\nabla_X^{1,0}Y) \otimes \omega + Y \otimes (\nabla_X^{0,1}\omega)$$

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

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

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

Example 2.10 Suppose that $T \in \Gamma(T_{0,2}M)$ (for example, $T = g$ or $T = \text{Ric}$). Then $T(X, Y) = C(X \otimes Y \otimes T)$ for a 2-fold contraction C , thus

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

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

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

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

Example 2.11 Let again D denote the Levi-Civita connection of $g = \langle \cdot, \cdot \rangle$. For a function $f \in C^\infty(M)$, the *gradient* $\text{grad } f \in \Gamma(TM)$ and the *Hesse form* $\text{Hess}(f) \in \Gamma(T_{0,2}M)$ are defined by the relations $\langle \text{grad } f, Y \rangle = df(Y)$ and

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

for $X, Y \in \Gamma(TM)$. By (3), this last term is equal to $(D_X^{0,1}df)(Y)$; briefly,

$$\text{Hess}(f) = D df = D^2 f.$$

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 $\text{Hess}(f)$ is symmetric. If c is a geodesic in M with $c(0) = p$ and $c'(0) = v$, then

$$(f \circ c)''(0) = \frac{d}{ds} \Big|_{s=0} \langle \text{grad } f(c(s)), c'(s) \rangle = \text{Hess}(f)_p(v, v).$$

Example 2.12 If R is the curvature tensor of a connection ∇ 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 2.10 that

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

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

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

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

Lemma 2.13 (second Bianchi identity) *Let R be the curvature tensor of a torsion-free connection ∇ . Then*

$$(\nabla_ZR)(X, Y)W + (\nabla_XR)(Y, Z)W + (\nabla_YR)(Z, X)W = 0$$

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 2.12 that

$$\begin{aligned}(\nabla_ZR)(X, Y)W &= \nabla_Z(R(X, Y)W) - R(X, Y)\nabla_ZW - R(\nabla_ZX, Y)W - R(X, \nabla_ZY)W \\ &= \nabla_Z[\nabla_X, \nabla_Y]W - [\nabla_X, \nabla_Y]\nabla_ZW - R(\nabla_ZX, Y)W + R(\nabla_ZY, X)W \\ &= [\nabla_Z, [\nabla_X, \nabla_Y]]W - R(\nabla_ZX, Y)W + R(\nabla_YZ, X)W,\end{aligned}$$

where in the last step we have used that $\nabla_ZY = \nabla_YZ$ since ∇ is torsion-free. Now it is easy to see that the cyclic sum in the lemma vanishes. $\square$

Lemma 2.14 *Let ∇ be a connection that is compatible with $g = \langle \cdot, \cdot \rangle$, and let $C = C^g : \Gamma(T_{0,2}M) \rightarrow C^\infty(M)$ denote the metric contraction. Then*

$$C(\nabla_X^{0,2}T) = X(C(T))$$

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 $c: (-\epsilon, \epsilon) \rightarrow M$ with $\dot{c}(0) = x$. Since ∇ is compatible with g , there exists a parallel orthonormal frame field $(E_1, \dots, E_m)$ along c . Then

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

by Example 2.10. 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 $DY: \Gamma(TM) \rightarrow \Gamma(TM)$. The *divergence* of a symmetric tensor field $T \in \Gamma(T_{0,2}M)$ is the metric contraction

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

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

Proposition 2.15 *For any Riemannian manifold M , $d \operatorname{scal} = 2 \operatorname{div}(\operatorname{Ric})$.*

Proof: Let $p \in M$, and let $v, w, x, y, z \in TM_p$. It follows from Lemma 2.13 and the symmetries of R (see Example 2.12) 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 $(e_1, \dots, e_m)$ 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 ,

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

Now it follows from Lemma 2.14 and the definition of the divergence that $d \operatorname{scal}(z) = z(\operatorname{scal}) = 2 \operatorname{div}(\operatorname{Ric})(z)$. $\square$

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

Theorem 2.16 (Schur 1886) *Let (M, g) be a connected Riemannian manifold of dimension $m \geq 3$.*

- (1) *If $\text{Ric} = fg$ for some $f \in C^\infty(M)$, then f is constant.*
- (2) *If for every $p \in M$ the sectional curvature $\text{sec}(P)$ is independent of $P \subset TM_p$, then sec is constant on M .*

A Riemannian manifold (M, g) with $\text{Ric} = \kappa g$ 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 also that $\text{Ric} = fg$ implies $\text{scal} = mf$.

Proof: By Proposition 2.15,

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

and using the relation $Dg = 0$ (see Example 2.10) one checks that $\text{div}(fg) = df$. Since $m \geq 3$, this yields $df = 0$. As M is connected, (1) follows.

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

2.4 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 $D_X Y = (\bar{D}_X Y)^\top$ 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 $TM_p^\perp$.

Definition 2.17 The $\mathbb{R}$ -bilinear map $h: \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(TM^\perp)$ defined by

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

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:

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

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

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

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

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

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 $N(p)$.

Lemma 2.18 *The operator $S_{N(p)}: TM_p \rightarrow TM_p$ is self-adjoint. For $x \in TM_p$,*

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

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 $F: (M, g) \rightarrow (\bar{M}, \bar{g})$. For this, one replaces $\bar{D}$ by the connection $\bar{D}^F$ along F and N by a unit normal vector field along F .

Theorem 2.19 (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

$$R(V, W, X, Y) = g(V, R(X, Y)W) = g(V, D_X D_Y W - D_Y D_X W).$$

With $D_X D_Y W = (\bar{D}_X D_Y W)^T = (\bar{D}_X \bar{D}_Y W - \bar{D}_X h(W, Y))^T$ and the corresponding expression for $D_Y D_X W$ we get that

$$\begin{aligned} R(V, W, X, Y) &= \bar{g}(V, \bar{R}(X, Y)W - \bar{D}_X h(W, Y) + \bar{D}_Y h(W, X)) \\ &= \bar{R}(V, W, X, Y) + \bar{g}(\bar{D}_X V, h(W, Y)) - \bar{g}(\bar{D}_Y V, h(W, X)), \end{aligned}$$

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

Lemma 2.20 *For a curve $c: I \rightarrow M \subset \bar{M}$ 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 $Y = \sum_i Y^i A_i \circ c$ for some smooth functions Y^i on I , and

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

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. $\square$

Proposition 2.21 (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 $c: I \rightarrow M$ 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 2.20 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 $c_v: (\alpha_v, \omega_v) \rightarrow M$ 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 2.20 that $h(v, v) = 0$ for all $v \in TM_p$, which is equivalent to (1) by the symmetry of h . $\square$

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 2.19 that $\sec(P)$ equals the Gauss curvature of the 2-sphere $L \cap S^m$ (which is 1, as we know).

2.5 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 $T\bar{M}_{\bar{p}} = TM_p \times TM'_{p'}$.

Definition 2.22 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 $\pi: \bar{M} \rightarrow M$ and $\pi': \bar{M} \rightarrow M'$ 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

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

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.

Proposition 2.23 (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 $P = \text{span}\{(v, 0), (0, w')\}$, $\sec_{\bar{M}}(P) = 0$.

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 $\sec_{\bar{M}}(P) = R(v, 0, v, 0) + R'(0, w', 0, w') = 0$. $\square$

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

3.1 Second variation of arc length

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

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

be a *piecewise smooth variation* of $c = \gamma_0$, that is, γ is continuous and there exists a finite subdivision $a = t_0 < t_1 < \dots < t_k = b$ of $[a, b]$ such that $\gamma|_{(-\epsilon, \epsilon) \times [t_{i-1}, t_i]}$ is C^∞ for $i = 1, \dots, k$. 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 $\frac{D}{dt} V_0$.

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

$$\left. \frac{d^2}{ds^2} \right|_{s=0} L(\gamma_s) = \int_a^b |(V'_0)^\perp|^2 - R(V_0, c', V_0, c') dt + \left\langle \left. \frac{D}{\partial s} \right|_{s=0} V, c' \right\rangle \Big|_a^b.$$

Here $(V'_0)^\perp$ denotes the part of V'_0 normal to c , thus

$$|(V'_0)^\perp|^2 = |V'_0|^2 - \langle V'_0, c' \rangle^2.$$

If the variation is itself *normal*, that is, if $\langle V_0, c' \rangle = 0$, then V'_0 is normal and so $(V'_0)^\perp = V'_0$. If the variation is *proper*, that is, if $\gamma_s(a) = c(a)$ and $\gamma_s(b) = c(b)$ for all s , then $\left. \frac{D}{\partial s} \right|_{s=0} V$ vanishes at $t = a$ and $t = b$, hence

$$\left. \frac{d^2}{ds^2} \right|_{s=0} L(\gamma_s) = \int_a^b |(V'_0)^\perp|^2 - R(V_0, c', V_0, c') dt.$$

For a proper and normal *smooth* variation, the formula may be rewritten as

$$\left. \frac{d^2}{ds^2} \right|_{s=0} L(\gamma_s) = - \int_a^b \langle V_0, V''_0 + R(V_0, c')c' \rangle dt,$$

because $|V'_0|^2 = \langle V_0, V'_0 \rangle' - \langle V_0, V''_0 \rangle$ and $\int_a^b \langle V_0, V'_0 \rangle' dt = 0$.

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

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

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

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

Furthermore, by Proposition 2.3 and the second part of Proposition 1.12,

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

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

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

Theorem 3.2 (Synge 1936) *Let M be a closed 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 $\alpha: [0, 1] \rightarrow M$ 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 $c: [0, l] \rightarrow M$ in the free homotopy class $[\alpha]$ of α , 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 $P: H \rightarrow H$, 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 $\gamma: (-\epsilon, \epsilon) \times [0, l] \rightarrow M$ 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 $\text{sec} > 0$ 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 $[\alpha]$. $\square$

Remark 3.3 Every non-orientable manifold M has a two-sheeted *orientable covering* $\pi: \bar{M} \rightarrow M$. 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 π 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.

3.2 Jacobi fields

Definition 3.4 A vector field Y along a geodesic $c: I \rightarrow M$ (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 $s \mapsto \tilde{c}(s) := c(\alpha s + \beta)$.

Lemma 3.5 *The set of Jacobi fields along a geodesic $c: I \rightarrow M^m$ 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 $(E_1, \dots, E_m)$ 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 $R(Y, c')c' = \sum_j Y^j R(E_j, c')c' = \sum_{i,j} Y^j \varrho_j^i E_i$ for some functions ϱ_j^i ; thus Y is a Jacobi field if and only if

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

for $i = 1, \dots, m$. This gives the result. $\square$

Proposition 3.6 *Let $c: [0, l] \rightarrow M$ be a unit speed geodesic. If γ is a variation of $c = \gamma_0: [0, l] \rightarrow M$ 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 γ_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 $\sigma: (-\epsilon, \epsilon) \rightarrow M$ with $\sigma(0) = c(0)$ and $\sigma'(0) = Y(0)$. Let X and W be any vector fields along σ such that $X(0) = c'(0)$ and $W(0) = Y'(0) - \frac{D}{ds}|_{s=0} X$. For $s \in (-\epsilon, \epsilon)$ sufficiently close to 0, the geodesic $\gamma_s: [0, l] \rightarrow M$ with

$$\gamma_s(0) = \sigma(s) \quad \text{and} \quad \gamma'_s(0) = T_0(s) := X(s) + sW(s)$$

is defined; note that $\gamma_0 = c$ since $T_0(0) = c'(0)$. 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 σ . Secondly,

$$V'_0(0) = \frac{D}{dt}\Big|_{t=0} V_0 = \frac{D}{ds}\Big|_{s=0} T_0 = \frac{D}{ds}\Big|_{s=0} X + W(0) = Y'(0)$$

by Proposition 1.12 and the choice of T_0 and W . □

Remark 3.7 If Y is a Jacobi field along the geodesic $c: I \rightarrow M$, 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

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

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

Example 3.8 (Jacobi fields in space forms) Let $c: \mathbb{R} \rightarrow M$ be a unit speed geodesic in a space form of curvature $\kappa \in \mathbb{R}$. Then $R(X, c')c' = \kappa X$ for any normal field $X \in \Gamma(c^*TM)$ along c . Hence, if E is a parallel unit normal field along c and $f: \mathbb{R} \rightarrow \mathbb{R}$ is a smooth function, then $Y = fE$ is a Jacobi field if and only if

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

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

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

Theorem 3.9 (Myers 1941) *Let (M, g) be a complete Riemannian manifold with $\text{Ric}(v, v) \geq (m-1)\kappa$ for all unit vectors $v \in TM$ and for some constant $\kappa > 0$. Then*

$$\text{diam}(M) := \sup\{d(p, q) : p, q \in M\} \leq \frac{\pi}{\sqrt{\kappa}},$$

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 $c: [0, l] \rightarrow M$ from p to q . Choose a parallel orthonormal frame field $(E_1, \dots, E_m)$ along c with $E_m = c'$. Put $\lambda := (\frac{\pi}{l})^2$, and note that the function $t \mapsto \text{sn}_\lambda(t) = \frac{1}{\sqrt{\lambda}} \sin(\sqrt{\lambda}t)$ vanishes at 0 and $l = \frac{\pi}{\sqrt{\lambda}}$ and is positive on $(0, l)$. For $i = 1, \dots, m-1$, consider a proper and normal variation of c with variation vector field $t \mapsto \text{sn}_\lambda(t) E_i(t)$ along c . Since c has minimal length $d(p, q)$, it follows from the second variation formula (Theorem 3.1) and the relation $\text{sn}_\lambda'' = -\lambda \text{sn}_\lambda$ that

$$\begin{aligned} 0 &\leq - \int_0^l \langle \text{sn}_\lambda E_i, \text{sn}_\lambda'' E_i + R(\text{sn}_\lambda E_i, c')c' \rangle dt \\ &= \int_0^l \text{sn}_\lambda^2 (\lambda - \langle E_i, R(E_i, c')c' \rangle) dt. \end{aligned}$$

By taking the sum from $i = 1$ to $m-1$ and using the bound on the Ricci curvature, we get that $0 \leq (m-1)(\lambda - \kappa)$, thus $l = \frac{\pi}{\sqrt{\lambda}} \leq \frac{\pi}{\sqrt{\kappa}}$. Since $p \neq q$ were arbitrary, this shows that $\text{diam}(M) \leq \frac{\pi}{\sqrt{\kappa}}$.

In particular, M is complete and bounded, hence compact. Furthermore, if $\pi: \tilde{M} \rightarrow M$ 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. $\square$

3.3 Conjugate points

Definition 3.10 Let $c: [a, b] \rightarrow M$ be a geodesic from p to q (where $a < b$, but not necessarily $p \neq q$). 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.

Remark 3.11 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 exists a unique Jacobi field Y along c with $Y(a) = v$ and $Y(b) = w$.

Lemma 3.12 *Let $c: [0, l] \rightarrow M$ 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 $Y'(0) = w \neq 0$. $\square$

Theorem 3.13 *Let $c: [0, l] \rightarrow M$ 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 $\gamma: [0, l] \rightarrow M$ 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 γ is a reparametrization of c .*

It is not true in general that $L(\gamma) \geq L(c)$ for all C^1 curves γ 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 3.12, $d(\exp_p)_{tv}: TM_p \rightarrow TM_{c(t)}$ is bijective for all $t \in [0, l]$. It follows that there exist $0 = t_0 < t_1 < \dots < t_k = l$ and open sets $U_1, \dots, U_k \subset TM_p$ 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 $i = 1, \dots, k$. Choose $\epsilon > 0$ such that every V_i contains the closed ϵ -neighborhood of $c([t_{i-1}, t_i])$. Now let $\gamma: [0, l] \rightarrow M$ 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 $\beta: [0, l] \rightarrow TM_p$ 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 β 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. $\square$

Definition 3.14 Let $c: [a, b] \rightarrow M$ 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

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

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

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

Lemma 3.15 *Let $c: [a, b] \rightarrow M$ 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 $a = t_0 < t_1 < \dots < t_k = b$ be such that $Y_i := Y|_{[t_{i-1}, t_i]}$ is smooth for $i = 1, \dots, k$. First, choose a smooth function f on $[a, b]$ that vanishes at $t_0, \dots, t_k$ and is positive elsewhere. The vector field X with $X|_{[t_{i-1}, t_i]} := f(Y_i'' + R(Y_i, c')c')$ is smooth, and

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

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

The following result complements Theorem 3.13.

Theorem 3.16 *Let $c: [0, l] \rightarrow M$ 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 $\gamma: (-\epsilon, \epsilon) \times [0, l] \rightarrow M$ of $c = \gamma_0$ such that $L(\gamma_s) < L(c)$ for all $s \in (-\epsilon, \epsilon) \setminus \{0\}$.*

Proof: Let $Y \neq 0$ 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 γ with (piecewise smooth) variation vector field $\lambda X + Y$ along c . $\square$

3.4 The Rauch comparison theorem

Proposition 3.17 (first index lemma) *Let $c: [0, l] \rightarrow M$ 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 3.11 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 3.13 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 3.15, 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 3.13. Thus $I(V_0, W) = 0$ for all such W , and so V_0 is a Jacobi field by the second part of Lemma 3.15. 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 [Ber1962] and a corollary.

Assumptions 3.18 Suppose that M and $\bar{M}$ are two Riemannian manifolds with $\dim(\bar{M}) \geq \dim(M) \geq 2$, and $c: [0, l] \rightarrow M$ and $\bar{c}: [0, l] \rightarrow \bar{M}$ 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 $\bar{c}'(t) \in \bar{P}$.

Theorem 3.19 (Rauch 1951) *Suppose, in addition to Assumptions 3.18, 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 $Y(0) = 0$, $\bar{Y}(0) = 0$, 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, $|Y^T| = |\bar{Y}^T|$ because $\langle Y, c' \rangle(0) = \langle \bar{Y}, \bar{c}' \rangle(0)$ and $\langle Y, c' \rangle'(0) = \langle \bar{Y}, \bar{c}' \rangle'(0)$.) We further assume that $\bar{Y} \neq 0$, 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 $(E_1, \dots, E_{m-1}, c')$ along c such that $\lambda E_1(r) = Y(r)$. Since $\dim(\bar{M}) \geq \dim(M)$, there exists an orthonormal system $(\bar{E}_1, \dots, \bar{E}_{m-1}, \bar{c}')$ 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 $\bar{c}$. Note that $|\bar{X}| = |Y|$ 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 $\bar{c}|_{[0, r]}$. The assumption involving the sectional curvatures yields the inequality

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

Furthermore, $\bar{I}^r(\bar{\lambda} \bar{X}, \bar{\lambda} \bar{X}) \geq \bar{I}^r(\lambda \bar{Y}, \lambda \bar{Y})$ by Proposition 3.17. We conclude that $\bar{\lambda}^2 I^r(Y, Y) \geq \lambda^2 \bar{I}^r(\bar{Y}, \bar{Y})$ or, in view of the second expression in Definition 3.14,

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

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

Corollary 3.20 *Let again $M, \bar{M}$ and $c, \bar{c}$ satisfy Assumptions 3.18 and suppose, as in Theorem 3.19, 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 $H: TM_p \rightarrow T\bar{M}_{\bar{p}}$ be a linear isometric embedding such that $H(v) = \bar{v}$. Then for all $w \in TM_p$ and $\bar{w} := H(w)$,*

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

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

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

thus $|Y(l)| \geq |\bar{Y}(l)|$ by Theorem 3.19. Since these Jacobi fields are given explicitly by $Y(t) = d(\exp_p)_{tv}(tw)$ and $\bar{Y}(t) = d(\exp_{\bar{p}})_{t\bar{v}}(t\bar{w})$ (see the proof of Lemma 3.12), the result follows. $\square$

Corollary 3.20 may further be used to compare distances, as shown next (see also Theorem 5.12), or volumes (see the proof of Theorem 6.1).

Corollary 3.21 *Let $M, \bar{M}$ be two Riemannian manifolds with $\dim(M) = \dim(\bar{M})$ and $\sec_M \leq \kappa \leq \sec_{\bar{M}}$ for some constant $\kappa \in \mathbb{R}$. Let $p \in M$ and $\bar{p} \in \bar{M}$, let $H: TM_p \rightarrow T\bar{M}_{\bar{p}}$ be a linear isometry, and suppose that $r > 0$ is such that both $B_r \subset TM_p$ and $B_r \subset T\bar{M}_{\bar{p}}$ are normal balls. Then the diffeomorphism*

$$F := \exp_{\bar{p}} \circ H \circ (\exp_p|_{B_r})^{-1}: B(p, r) \rightarrow B(\bar{p}, r)$$

does not increase the length of any piecewise C^1 curve in $B(p, r)$, in particular $d(F(x), F(y)) \leq d(x, y)$ for all $x, y \in B(p, \frac{r}{2})$.

Proof: Note that since $B_r \subset T\bar{M}_{\bar{p}}$ is normal, there is no point conjugate to $\bar{p}$ along a radial segment in $B(\bar{p}, r)$. It follows from Corollary 3.20 that if $c: [a, b] \rightarrow B(p, r)$ is any piecewise C^1 curve, and $\gamma := (\exp_p|_{B_r})^{-1} \circ c$, then

$$\int_a^b |d(\exp_{\bar{p}})_{H(\gamma(t))} \circ H(\gamma'(t))| dt \leq \int_a^b |d(\exp_p)_{\gamma(t)}(\gamma'(t))| dt,$$

thus $L(F \circ c) = \int_a^b |(F \circ c)'(t)| dt \leq \int_a^b |c'(t)| dt = L(c)$. If $x, y \in B(p, \frac{r}{2})$, then $d(x, y) < r$ and every curve c of length $< r$ from x to y lies in $B(p, r)$, hence $L(F \circ c) \leq L(c)$, and by taking the infimum over all such c we conclude that $d(F(x), F(y)) \leq d(x, y)$. $\square$

3.5 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 $c: [0, l] \rightarrow M$ is a unit speed geodesic from $p \in \Sigma$ to $q \in M$ with $c'(0) = N(p)$. Recall that the shape operator of (Σ, N) at p is the linear operator $S_N = S_{N(p)}: T\Sigma_p \rightarrow T\Sigma_p$ satisfying $S_N(x) = -(D_x N)^\top$ (Lemma 2.18).

Proposition 3.22 *For a Jacobi field Y along c , the following are equivalent:*

- (1) Y is the variation vector field along c of a variation γ of $c = \gamma_0$ such that every curve $\gamma_s = \gamma(s, \cdot)$ is a geodesic with $\gamma_s(0) \in \Sigma$ and $\gamma'_s(0) \in T\Sigma_{\gamma_s(0)}^\perp$.
- (2) $Y(0) \in T\Sigma_p$ and $Y'(0)^T = (D_{Y(0)}N)^T = -S_N(Y(0))$.

A Jacobi field with these properties is called a Σ -Jacobi field along c .

Proof: Suppose first that (1) holds. Put $\sigma(s) := \gamma_s(0) \in \Sigma$ and $T_0(s) := \gamma'_s(0) \in T\Sigma_{\sigma(s)}^\perp$. Then $Y(0) = \sigma'(0) \in T\Sigma_p$. Moreover,

$$Y'(0) = \frac{D}{dt}\Big|_{t=0} Y = \frac{D}{ds}\Big|_{s=0} T_0.$$

Note that $T_0(0) = c'(0) = N(p) = (N \circ \sigma)(0)$. Hence, if Z is any vector field along σ tangent to Σ , then $\langle \frac{D}{ds}(T_0 - N \circ \sigma), Z \rangle = -\langle T_0 - N \circ \sigma, \frac{D}{ds}Z \rangle$ vanishes at $s = 0$. It follows that

$$\left(\frac{D}{ds}\Big|_{s=0} T_0\right)^T = \left(\frac{D}{ds}\Big|_{s=0} (N \circ \sigma)\right)^T = (D_{Y(0)}N)^T.$$

Combining these equalities we get (2).

Conversely, suppose that (2) holds. Choose a curve $\sigma: (-\epsilon, \epsilon) \rightarrow \Sigma$ with $\sigma'(0) = Y(0) \in T\Sigma_p$. Let X and W be any vector fields along σ normal to Σ such that $X(0) = N(p)$ and $W(0)$ equals the normal component of $Y'(0) - \frac{D}{ds}\Big|_{s=0} X$. Put $T_0(s) := X(s) + sW(s)$. Note that $T_0(0) = N(p) = c'(0)$. Similarly as above, $(\frac{D}{ds}\Big|_{s=0} X)^T = (D_{Y(0)}N)^T$, thus

$$\frac{D}{ds}\Big|_{s=0} T_0 = \frac{D}{ds}\Big|_{s=0} X + W(0) = (D_{Y(0)}N)^T + Y'(0)^\perp = Y'(0)$$

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

Definition 3.23 The point $q = c(l)$ is called a *focal point* of Σ along c if there exists a non-trivial Σ -Jacobi field Y along c with $Y(l) = 0$.

In analogy to Lemma 3.12 it can be shown that q is focal point of Σ along c if and only if $lc'(0)$ is a singular point of the *normal exponential map* $\exp^\perp := \exp|_{\Omega \cap T\Sigma^\perp}$.

Remark 3.24 Suppose now that $\dim(\Sigma) = \dim(M) - 1$ and $S_{c'(0)} = 0$. (This holds, for example, if Σ 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 3.22 that a normal Jacobi field Y along c is a Σ -Jacobi field if and only if $Y'(0) = 0$. Hence, $q = c(l)$ is a focal point of Σ 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 by Remark 3.7. If no such Jacobi field 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 .

Proposition 3.25 (second index lemma) *Let $c: [0, l] \rightarrow M$ 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 $i = 1, \dots, m-1$. 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 $\lambda_1, \dots, \lambda_{m-1}: [0, l] \rightarrow \mathbb{R}$ 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 $Y'(0) = 0$,

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

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 $f_i = \sum_j \lambda_j' \langle V_j, V_i \rangle + \lambda_j (\langle V_j', V_i \rangle - \langle V_j, V_i' \rangle)$, 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

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

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

Theorem 3.26 (Berger 1962) *Let again $M, \bar{M}$ and $c, \bar{c}$ be given as in Assumptions 3.18, 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 $Y'(0) = 0, \bar{Y}'(0) = 0$, 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 $t \in [0, l]$.

Proof: Suppose, as in the proof of Theorem 3.19, that $Y, \bar{Y}$ are normal along $c, \bar{c}$ and $\bar{Y}$ is non-trivial. Then it follows from the assumptions that $\bar{Y}(t) \neq 0$ for all $t \in [0, l]$. Now the same argument as in the proof of Theorem 3.19, with Proposition 3.25 in place of Proposition 3.17, shows again that $(|Y|^2/|\bar{Y}|^2)' \geq 0$ on $(0, l)$. Since $|Y(0)| = |\bar{Y}(0)| \neq 0$, this yields the result. $\square$

Chapter 4

Riemannian submersions and coverings

4.1 Riemannian submersions

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

$$d\pi_p|_{H_p}: H_p \rightarrow TM_{\pi(p)}$$

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 $\dim(\bar{M}) - \dim(M)$.

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

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

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

Proposition 4.3 Suppose that $\pi: \bar{M} \rightarrow M$ is a Riemannian submersion.

- (1) Let c be a geodesic in M with $c(0) = \pi(p)$ for some $p \in \bar{M}$. Then there exists a unique maximal horizontal curve $\bar{c}: I \rightarrow \bar{M}$ 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 π 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, π 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 an integral curve of X with $\pi \circ \bar{c}(0) = c(0)$, and so $\pi \circ \bar{c} = c$ on some subinterval of $(-\epsilon, \epsilon)$. 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 $\bar{c}: I \rightarrow \bar{M}$, let γ 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 $c := \pi \circ \bar{c} = \pi \circ \bar{\gamma}|_I = \gamma|_I$. $\square$

Example 4.4 (warped products) Let (M, g) and (M', g') be Riemannian manifolds, and let $f: M \rightarrow (0, \infty)$ be a C^∞ 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

$$\pi: (M \times M', g \times_f g') \rightarrow (M, g)$$

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 2.21).

Proposition 4.5 *Let $\pi: \bar{M} \rightarrow M$ 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 $h \in \text{Isom}(\bar{M}, \bar{g})$ such that $h(p) = q$ and $\pi \circ h = \pi$. Then there exists a unique Riemannian metric g on M such that $\pi: (\bar{M}, \bar{g}) \rightarrow (M, g)$ 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

$$(d\pi_q|_{H_q})^{-1} \circ (d\pi_p|_{H_p}): H_p \rightarrow H_q \quad (1)$$

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 π is surjective, it now follows that for every $x \in M$ there is a unique inner product g_x on TM_x such that $d\pi_p|_{H_p}: H_p \rightarrow TM_x$ is an isometry for all $p \in \pi^{-1}\{x\}$. It can be shown that g_x depends smoothly on x . $\square$

Example 4.6 (Hopf fibration) 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

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

Then the quotient projection $\pi: S^{2n+1} \rightarrow \mathbb{C}P^n$ is a surjective submersion, the (generalized) Hopf fibration. For any λ as above, the map $h_\lambda: S^{2n+1} \rightarrow S^{2n+1}$ sending z to λz (a rotation) is an isometry of $(S^{2n+1}, g^{\text{sph}})$ satisfying $\pi \circ h_\lambda = \pi$. By Proposition 4.5 there exists a unique Riemannian metric g on $\mathbb{C}P^n$ with respect to which π is a Riemannian submersion, called the *Fubini–Study metric*.

For $p \in S^{2n+1}$ and a unit vector $v \in TS_p^{2n+1}$, let $\bar{c} = \bar{c}_v: \mathbb{R} \rightarrow S^{2n+1}$ be the geodesic given by $\bar{c}(t) = p \cos(t) + v \sin(t)$. Note that the fiber of π through p is contained in the plane with orthonormal basis $\{p, ip\}$, thus the vector $ip \in TS_p^{2n+1}$ 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 π .

4.2 Curvature of Riemannian submersions

Let $\pi: \bar{M} \rightarrow M$ 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 $\bar{X} \in \Gamma(T\bar{M})$ denote the unique horizontal vector field such that $\pi_*\bar{X} = X \circ \pi$ (that is, $\bar{X}$ and X are π -related). Notice that for all $X, Y \in \Gamma(TM)$,

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

Proposition 4.7 *For all $X, Y, Z \in \Gamma(TM)$,*

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

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

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

and

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

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

$$(\bar{D}_{\bar{X}}\bar{Y})^{\text{hor}} = \overline{D_X Y}. \quad (1)$$

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

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

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

$$(\bar{D}_{\bar{X}}\bar{Y})^{\text{ver}} = \frac{1}{2}[\bar{X}, \bar{Y}]^{\text{ver}}, \quad (2)$$

and (1) and (2) yield the second assertion. $\square$

Theorem 4.8 (O'Neill 1966) *For all $X, Y \in \Gamma(TM)$,*

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

Proof: Let $V, W, X, Y \in \Gamma(TM)$. From Proposition 4.7 we get that

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

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

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

It follows that

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

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

$$\langle \bar{V}, \bar{D}_{\bar{Y}} \bar{D}_{\bar{X}} \bar{W} \rangle = \langle V, D_Y D_X W \rangle \circ \pi - B.\tag{2}$$

Next, note that $\pi_* [\bar{X}, \bar{Y}]^{\text{hor}} = \pi_* [\bar{X}, \bar{Y}] = [X, Y] \circ \pi$ and thus $[\bar{X}, \bar{Y}]^{\text{hor}} = \overline{[X, Y]}$. Write $[\bar{X}, \bar{Y}] = \overline{[X, Y]} + T$ for $T := [\bar{X}, \bar{Y}]^{\text{ver}}$. Since T is vertical, so is $[\bar{W}, T]$. From Proposition 4.7 we get that

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

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

$$\langle \bar{V}, \bar{D}_{[\bar{X}, \bar{Y}]} \bar{W} \rangle = \langle \bar{V}, \overline{D_{[X, Y]} W} \rangle + C = \langle V, D_{[X, Y]} W \rangle \circ \pi + C.\tag{3}$$

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

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

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

Example 4.9 (curvature of complex projective space) Consider again the Hopf fibration $\pi: S^{2n+1} \rightarrow \mathbb{C}P^n$, where $\mathbb{C}P^n$ is equipped with the Fubini–Study metric (see Example 4.6). 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 $\bar{X}, \bar{Y} \in \Gamma(TS^{2n+1})$,

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

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 $\langle \bar{X}, \bar{D}_{\bar{Y}} T \rangle = \langle \bar{X}, i\bar{Y} \rangle$. 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 4.8 gives the explicit formula

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

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

4.3 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 $F: \bar{M} \rightarrow M$ 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 $dF_p: T\bar{M}_p \rightarrow TM_{F(p)}$ is bijective for every $p \in \bar{M}$. Given Riemannian metrics $\bar{g}$ and g on $\bar{M}$ and M , respectively, a smooth map $F: \bar{M} \rightarrow M$ 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).

Lemma 4.10 *Suppose that $F, G: \bar{M} \rightarrow M$ 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 $\bar{M}$. $\square$

Next, recall that a *covering map* $\pi: \bar{M} \rightarrow M$ 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 π .

Definition 4.11 Let again $(\bar{M}, \bar{g})$ and (M, g) be two Riemannian manifolds of the same dimension. A smooth covering map $\pi: \bar{M} \rightarrow M$ with the property that $\pi^*g = \bar{g}$ is called a *Riemannian covering map*.

Proposition 4.12 *Suppose that $\bar{M}$ is a complete Riemannian manifold and M is a connected Riemannian manifold of the same dimension. Then every local isometry $F: \bar{M} \rightarrow M$ 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 $\exp_q: B_r \rightarrow B(q, r)$ 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 $F|_{B(p, r)}: B(p, r) \rightarrow B(q, r)$ 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 $dF_p: TM_p \rightarrow TM_q$ 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 exists 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 $\Gamma \subset \text{Homeo}(\bar{M})$ of homeomorphisms of $\bar{M}$ acts *freely* on $\bar{M}$ if $\gamma(p) \neq p$ whenever $\gamma \in \Gamma \setminus \{\text{id}\}$ and $p \in \bar{M}$, and Γ 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 $\pi: \bar{M} \rightarrow \bar{M}/\Gamma$ is a covering map. Conversely, suppose that $F: \bar{M} \rightarrow M$ is a covering map onto another topological manifold M (it suffices to assume that M is a Hausdorff space). Then a homeomorphism $\gamma \in \text{Homeo}(\bar{M})$ is called a *deck transformation* or *covering transformation* if $F \circ \gamma = F$, and the group Γ 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 Γ is isomorphic to the fundamental group $\pi_1(M)$.

Let now again $\bar{M}$ be a connected *smooth* manifold, and let $\Gamma \subset \text{Diff}(\bar{M})$ be a group of (C^∞) diffeomorphisms of $\bar{M}$ that acts freely and properly discontinuously. Then there is a unique C^∞ structure on $\bar{M}/\Gamma$ such that the covering map $\pi: \bar{M} \rightarrow \bar{M}/\Gamma$ is a local diffeomorphism. Furthermore, if $\Gamma \subset \text{Isom}(\bar{M})$ 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 π is a local isometry (compare Proposition 4.5).

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 $\mathbb{M}_\kappa^m$ denote the

m -dimensional model space of constant sectional curvature κ ,

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

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

Theorem 4.13 (Killing 1891, Hopf 1926) *Let M be an m -dimensional space form of curvature $\kappa \in \mathbb{R}$. Then there exists a group $\Gamma \subset \text{Isom}(\mathbb{M}_\kappa^m)$ that acts freely and properly discontinuously on $\mathbb{M}_\kappa^m$ such that M is isometric to $\mathbb{M}_\kappa^m/\Gamma$. If M is simply connected, then M is isometric to $\mathbb{M}_\kappa^m$.*

Proof: Choose points $p \in \mathbb{M}_\kappa^m$ and $q \in M$, a linear isometry $H: T(\mathbb{M}_\kappa^m)_p \rightarrow TM_q$, and define

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

Since both $\mathbb{M}_\kappa^m$ and M have constant curvature κ , it follows from Corollary 3.20 that F is a local isometry; the assumption excluding conjugate points is satisfied in view of the explicit description of normal Jacobi fields in space forms, see Example 3.8.

Consider first the case $\kappa \leq 0$. Then F is defined on all of $\mathbb{M}_\kappa^m$, and $F: \mathbb{M}_\kappa^m \rightarrow M$ is a Riemannian covering map according to Proposition 4.12. The group $\Gamma \subset \text{Homeo}(\mathbb{M}_\kappa^m)$ of deck transformations of F acts freely and properly discontinuously on $\mathbb{M}_\kappa^m$. In fact, $\Gamma \subset \text{Diff}(\mathbb{M}_\kappa^m)$, because F is a local diffeomorphism, and there is an induced C^∞ structure on $\mathbb{M}_\kappa^m/\Gamma$ such that the projection $\pi: \mathbb{M}_\kappa^m \rightarrow \mathbb{M}_\kappa^m/\Gamma$ is a local diffeomorphism and $\mathbb{M}_\kappa^m/\Gamma$ is diffeomorphic to M . For $\gamma \in \Gamma$ and $p' \in \mathbb{M}_\kappa^m$, 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 $\Gamma \subset \text{Isom}(\mathbb{M}_\kappa^m)$, and $\mathbb{M}_\kappa^m/\Gamma$ carries a Riemannian metric such that π is a local isometry. Thus M is isometric to $\mathbb{M}_\kappa^m/\Gamma$.

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

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

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 4.10 shows that F and $\tilde{F}$ agree on the intersection of their domains, $\mathbb{M}_\kappa^m \setminus \{-p, -\tilde{p}\}$, and therefore F extends to a local isometry $F: \mathbb{M}_\kappa^m \rightarrow M$ with $F(-p) = \tilde{F}(-p)$. The rest of the argument is the same as in the case $\kappa \leq 0$. $\square$

Example 4.14 (Euclidean space forms) Let $\Gamma \subset \text{Isom}(\mathbb{R}^m, g^{\text{eucl}})$ be a group of translations that acts freely and properly discontinuously. Then there exist linearly independent vectors $v_1, \dots, v_k \in \mathbb{R}^m$ such that Γ is the group

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

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

If $\Gamma \subset \text{Isom}(\mathbb{R}^m, g^{\text{eucl}})$ acts freely and properly discontinuously with compact quotient, then the group $\Gamma' \subset \Gamma$ of translations has finite index in Γ and $\mathbb{R}^m/\Gamma'$ is finitely covered by the flat torus $\mathbb{R}^m/\Gamma'$. This is due to Bieberbach [Bi1911], [Bi1912]. See [Buse1985] for an elegant geometric proof.

Example 4.15 (hyperbolic space forms) Every compact oriented surface of genus $n \geq 2$ can be realized, in a flexible way, as a quotient $(H^2, g^{\text{hyp}})/\Gamma$. 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^{hyp}) with $m \geq 3$ the famous *Mostow rigidity theorem* holds [Mos1968], [Mos1973]: if H^m/Γ and H^m/Γ' are compact and Γ, Γ' are isomorphic, then H^m/Γ and H^m/Γ' are isometric. This holds more generally for quotients of irreducible symmetric spaces of non-compact type and dimension at least three. A proof of this result in the (rank one) case of negative curvature will be given at the end of Chapter 8.

Example 4.16 (spherical space forms) For the standard sphere (S^m, g^{sph}) and $\Gamma = \{\text{id}, -\text{id}\}$,

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

is the real projective space. The canonical (quotient) metric g^{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 4.17 below.

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

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

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

Theorem 4.17 *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 $\mathbb{R}P^m = S^m/\{\text{id}, -\text{id}\}$.*

Proof: From Theorem 4.13, we know that $M = S^m/\Gamma$ for some group $\Gamma \subset \text{Isom}(S^m) \simeq O(m+1)$ that acts freely and properly discontinuously on S^m . Let $\gamma \in \Gamma$. Since $m+1$ is odd, γ has an eigenvalue 1 or -1 . If γ has an eigenvalue 1, then $\gamma = \text{id}$ because Γ acts freely. If -1 is the only real eigenvalue of γ , then 1 is an eigenvalue of γ^2 and hence $\gamma^2 = \text{id}$. But then $\gamma = -\text{id}$, 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 $\Gamma = \{\text{id}\}$ or $\Gamma = \{\text{id}, -\text{id}\}$. $\square$

A general reference for space forms is [Wo2011].

4.4 Hadamard manifolds

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

Theorem 4.18 (Hadamard 1898, Cartan 1928) *Let M be a complete Riemannian manifold with sectional curvature $\text{sec} \leq 0$, and let $p \in M$. Then $\exp_p: TM_p \rightarrow M$ 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 $\text{sec} \leq 0$, geodesics in (M, g) have no conjugate points (exercise). Hence, by Lemma 3.12, $\exp_p: TM_p \rightarrow M$ 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 4.12 shows that $\exp_p: TM_p \rightarrow M$ is a covering map. $\square$

Lemma 4.19 *Let M be a Hadamard manifold. Then for every pair of geodesics $c, \bar{c}: \mathbb{R} \rightarrow M$, the function $s \mapsto h(s) := d(c(s), \bar{c}(s))$ is convex on $\mathbb{R}$.*

Proof: Fix $r \in \mathbb{R}$ and put $l := h(r)$. It follows from Theorem 4.18 that for every $s \in \mathbb{R}$ there exists a *unique* geodesic $\gamma_s: [0, l] \rightarrow M$ from $c(s)$ to $\bar{c}(s)$, and $L(\gamma_s) = h(s) = d(c(s), \bar{c}(s))$. This yields a smooth variation $\gamma: \mathbb{R} \times [0, l] \rightarrow M$ of the unit speed geodesic γ_r . The variation vector field along γ_s satisfies $V_s(0) = c'(s)$ and $V_s(l) = \bar{c}'(s)$, in particular $\frac{D}{ds}\big|_{s=r} V_s(t)$ vanishes for $t = 0$ and $t = l$. Hence, the second variation formula (Theorem 3.1) gives

$$h''(r) = \int_0^l |(V_r')^\perp|^2 - R(V_r, \gamma_r', V_r, \gamma_r') dt.$$

Since M has non-positive curvature, $h''(r) \geq 0$. $\square$

Proposition 4.20 (flat strip) *Suppose that M is a Hadamard manifold, $c, \bar{c}: \mathbb{R} \rightarrow M$ are two unit speed geodesics with distinct images, and $\sup_{s \in \mathbb{R}} d(c(s), \bar{c}(s)) < \infty$. Let $f: \mathbb{R} \times [0, 1] \rightarrow M$ 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 exists 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 4.19.

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 $\delta \geq 0$,

$$d(f(x), f(\bar{x})) = \begin{cases} \delta v\left(\frac{\bar{s}-s}{\delta}\right) & \text{if } \delta > 0, \\ |\bar{s} - s| & \text{if } \delta = 0. \end{cases} \quad (1)$$

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

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

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

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

Similarly, we infer that

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

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

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

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

$$d(f(x), f(\bar{x})) = \lim_{\delta \rightarrow 0+} d(f(x), f(\bar{x}_\delta)) = \lim_{\delta \rightarrow 0+} \delta v\left(\frac{\bar{s}-s}{\delta}\right).$$

Since $|r| - v(0) \leq v(r) \leq |r| + v(0)$ 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 $\|\cdot\|$. $\square$

4.5 Isometries of Hadamard manifolds

Definition 4.21 Let M be a Hadamard manifold, and let $\gamma \in \text{Isom}(M)$ be an isometry. The *displacement function* $d_\gamma: M \rightarrow [0, \infty)$ is defined by $d_\gamma(p) := d(p, \gamma(p))$ for all $p \in M$. We put

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

Lemma 4.22 *The set $\text{Min}(\gamma)$ is closed, γ -invariant, and convex.*

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

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

thus $\text{Min}(\gamma)$ is γ -invariant. Lastly, if $\lambda \mapsto p_\lambda$ is a geodesic $[0, 1] \rightarrow M$ with endpoints $p_0, p_1 \in \text{Min}(\gamma)$, then it follows from Lemma 4.19 that

$$d(p_\lambda, \gamma(p_\lambda)) \leq (1 - \lambda) d(p_0, \gamma(p_0)) + \lambda d(p_1, \gamma(p_1)) = |\gamma|$$

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

Isometries of Hadamard manifolds are classified as follows.

Definition 4.23 An isometry γ of M is called *parabolic* if $\text{Min}(\gamma)$ is empty and *semi-simple* otherwise. In the latter case, γ is *elliptic* if $|\gamma| = 0$ (that is, γ has a fixed point) and *hyperbolic* if $|\gamma| > 0$.

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

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

If γ possesses an axis, then γ is called *axial*.

Lemma 4.24 *Let $\gamma \in \text{Isom}(M)$. If γ has an axis c , then the corresponding number $a > 0$ is equal to $|\gamma|$, thus γ is hyperbolic and $c(\mathbb{R}) \subset \text{Min}(\gamma)$. Conversely, if γ is hyperbolic, then for every point $p \in \text{Min}(\gamma)$ there exists an axis of γ through p , and for every pair of axes $c, \bar{c}: \mathbb{R} \rightarrow M$ the function $s \mapsto d(c(s), \bar{c}(s))$ is constant.*

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

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

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

For the converse, let γ be a hyperbolic isometry of M , put $a := |\gamma| > 0$, and let $p \in \text{Min}(\gamma)$. Let $c: [0, a] \rightarrow M$ be the (unit speed) geodesic from p to $\gamma(p)$, and extend it to a curve $c: \mathbb{R} \rightarrow M$ 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 $\bar{c}: \mathbb{R} \rightarrow M$ is another axis of γ , then the function $s \mapsto d(c(s), \bar{c}(s))$ is bounded and convex, hence constant. $\square$

Proposition 4.25 *Suppose that M is Hadamard manifold, and Γ is a subgroup of $\text{Isom}(M)$ 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 Γ is also free, then every $\gamma \in \Gamma \setminus \{\text{id}\}$ 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 $d_\gamma(p_k) \rightarrow |\gamma|$ for $k \rightarrow \infty$. There exist a compact set $K \subset M$ and elements $\alpha_k \in \Gamma$ with $q_k := \alpha_k^{-1}(p_k) \in K$ 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 Γ 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 $i \rightarrow \infty$. Then $d_\gamma(\alpha_{k(i)}(q)) = d(q, \gamma_{k(i)}(q)) = d_{\bar{\gamma}}(q)$ for all i , and

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

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

The following result was established in [Pr1942].

Theorem 4.26 (Preissmann 1942) *If M is a closed Riemannian manifold with $\text{sec} < 0$, 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 $\text{sec} < 0$.

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 $\Gamma \subset \text{Isom}(\tilde{M})$ of deck transformations, which is isomorphic to $\pi_1(M)$. The group Γ acts freely and properly discontinuously as well as cocompactly on $\tilde{M}$, because M is compact. By Proposition 4.25, every $\gamma \in \Gamma \setminus \{\text{id}\}$ is axial. Since $\sec < 0$, it follows from Proposition 4.20 (flat strip) that every such γ has a unique axis $L_\gamma \subset \tilde{M}$. Furthermore, since γ has no fixed point, every line $L \subset \tilde{M}$ preserved by γ is an axis and hence equal to L_γ . Now if $\beta, \gamma \in \Gamma \setminus \{\text{id}\}$ 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 $\mathbb{Z}$. $\square$

As observed in [By1970] (see also pp. 263–264 in [DoC1992]), Theorem 4.26 holds more generally for solvable subgroups; moreover, $\pi_1(M)$ contains no solvable subgroup of finite index. A further generalization is the Flat Torus Theorem, proved independently by Gromoll–Wolf [GroW1971] and Lawson–Yau [LawY1972]: if M is a closed Riemannian manifold of non-positive curvature, then $\pi_1(M)$ contains a free abelian subgroup of rank k if and only if there exists a flat k -torus isometrically and totally geodesically immersed in M .

Chapter 5

Triangle comparison

5.1 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 $[0, d(p, q)] \rightarrow M$ 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

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

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 $\text{Per}(H)$ of a hinge $H = H_p(x, y)$ we mean the perimeter of the triple (p, x, y) .

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

$$D_\kappa := \text{diam}(\mathbb{M}_\kappa^m) = \begin{cases} \frac{\pi}{\sqrt{\kappa}} & \text{if } \kappa > 0, \\ \infty & \text{if } \kappa \leq 0. \end{cases}$$

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

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

(Compare Example 3.8.) Explicitly,

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

Note that sn_κ and $\operatorname{sn}'_\kappa = \operatorname{cs}_\kappa$ are positive on $(0, D_\kappa)$ and $(-\frac{1}{2}D_\kappa, \frac{1}{2}D_\kappa)$, respectively. The identity $\operatorname{cs}_\kappa^2 + \kappa \operatorname{sn}_\kappa^2 = 1$ holds, and for $x, y \in \mathbb{R}$,

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

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

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

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

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

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

for $\kappa \neq 0$.

Proof: See, for example, Chapter I.2 in [BrH1999]. □

Lemma 5.2 *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 $\angle_p(x, y) = \gamma$, 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 [Alex1955] (compare Lemma 4.3.3 in [BuBI2001]).

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

$$\angle_q(p, y) + \angle_q(x, y) \leq \pi \iff \angle_p(q, y) \geq \angle_{\bar{p}}(\bar{x}, \bar{y}),$$

$$\angle_q(p, y) + \angle_q(x, y) \geq \pi \iff \angle_p(q, y) \leq \angle_{\bar{p}}(\bar{x}, \bar{y}).$$

Proof: Prolongate pq to a segment px' of length $|px'| = |pq| + |qx|$. Consider the following obvious identities:

$$\pi - \angle_q(p, y) - \angle_q(x, y) = \angle_q(x', y) - \angle_q(x, y), \quad (1)$$

$$|x'y| - |xy| = |x'y| - |\bar{x}\bar{y}|, \quad (2)$$

$$\angle_p(x', y) - \angle_{\bar{p}}(\bar{x}, \bar{y}) = \angle_p(q, y) - \angle_{\bar{p}}(\bar{x}, \bar{y}). \quad (3)$$

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). $\square$

5.2 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 $\mathbb{M}_\kappa^2$ is called a *comparison triple* for (p, x, y) if $|\bar{p}\bar{x}| = |px|$, $|\bar{p}\bar{y}| = |py|$, and $|\bar{x}\bar{y}| = |xy|$.

Remark 5.4 If $\kappa \leq 0$, such a comparison triple always exists, and if $\kappa > 0$, a comparison triple exists if and only if $\text{Per}(p, x, y) \leq 2D_\kappa$. This is obvious if one of the distances $a := |py|$, $b := |px|$, and $c := |xy|$ is zero or equal to D_κ . 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 $a + b \geq D_\kappa$. In either case, the given number c is contained in I , so there exists a unique $\gamma \in [0, \pi]$ such that $c_{a,b}(\gamma) = c$.

Definition 5.5 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 $\text{Per}(p, x, y) \leq 2D_\kappa$. Then any comparison triple $(\bar{p}, \bar{x}, \bar{y})$ in $\mathbb{M}_\kappa^2$ uniquely determines a hinge $H_{\bar{p}}(\bar{x}, \bar{y})$ and one defines the *comparison angle* $\angle_p^\kappa(x, y) \in [0, \pi]$ as the hinge angle, thus

$$\angle_p^\kappa(x, y) := \angle_{\bar{p}}(\bar{x}, \bar{y}).$$

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

$$\angle_p(x, y) := \limsup_{\substack{u \in px, v \in py \\ u, v \rightarrow p}} \angle_p^\kappa(u, v).$$

Remark 5.6 It is not difficult to see that the upper angle $\angle_p(x, y)$ does not depend on the choice of $\kappa \in \mathbb{R}$. Furthermore, for any three non-degenerate segments px, py, pz in M , the triangle inequality

$$\angle_p(x, y) + \angle_p(y, z) \geq \angle_p(x, z)$$

holds, see [Alex1951] or Proposition I.1.14 in [BrH1999].

Definition 5.7 Let again $H_p(x, y)$ be a hinge in M , and suppose that $\text{Per}(p, x, y) < 2D_\kappa$. Let $(\bar{p}, \bar{x}, \bar{y})$ be a comparison triple in $\mathbb{M}_\kappa^2$ for (p, x, y) , and let $H_{\bar{p}}(\hat{x}, \hat{y})$ be a *comparison hinge* in $\mathbb{M}_\kappa^2$ for $H_p(x, y)$; that is, $|\hat{p}\hat{x}| = |px|$, $|\hat{p}\hat{y}| = |py|$, and $\angle_{\bar{p}}(\hat{x}, \hat{y}) = \angle_p(x, y)$. We are interested in the following *comparison properties for curvature $\geq \kappa$* that $H_p(x, y)$ may or may not have:

(A $_\kappa$) (Angle comparison) $\angle_p(x, y) \geq \angle_{\bar{p}}(\bar{x}, \bar{y}) (= \angle_{\bar{p}}^\kappa(x, y))$;

(H $_\kappa$) (Hinge comparison) $|xy| \leq |\hat{x}\hat{y}|$;

(C $_\kappa$) (Chord 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 (A $^\kappa$), (H $^\kappa$), and (C $^\kappa$), are defined analogously, just with reversed inequalities.

Lemma 5.8 For an individual hinge $H_p(x, y)$ in M with $\text{Per}(p, x, y) < 2D_\kappa$,

$$(C_\kappa) \Rightarrow (A_\kappa) \Leftrightarrow (H_\kappa) \quad \text{and} \quad (C^\kappa) \Rightarrow (A^\kappa) \Leftrightarrow (H^\kappa).$$

For the implications $(A_\kappa) \Rightarrow (C_\kappa)$ and $(A^\kappa) \Rightarrow (C^\kappa)$, see Lemma 5.9 below.

Proof: Suppose that $H_p(x, y)$ satisfies (C $_\kappa$). With the above notation, it follows from the inequality $|uv| \geq |\bar{u}\bar{v}|$ and Lemma 5.2 that $\angle_p^\kappa(u, v) \geq \angle_{\bar{p}}^\kappa(\bar{u}, \bar{v}) = \angle_{\bar{p}}(\bar{x}, \bar{y})$. Taking the upper limit for $u, v \rightarrow p$, we conclude that (A $_\kappa$) holds. Lemma 5.2 also shows that $\angle_p(x, y) = \angle_{\bar{p}}(\hat{x}, \hat{y}) \geq \angle_{\bar{p}}(\bar{x}, \bar{y})$ if and only if $|\hat{x}\hat{y}| \geq |\bar{x}\bar{y}| = |xy|$; that is, (A $_\kappa$) $\Leftrightarrow$ (H $_\kappa$). The second part of the lemma is shown analogously. $\square$

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 $\angle_q(p, y) + \angle_q(x, y) = \pi$. Note that, by the triangle inequality for angles (Remark 5.6), the inequality $\angle_q(p, y) + \angle_q(x, y) \geq \pi$ always holds, as $\angle_q(p, x) = \pi$. Evidently, in a Riemannian manifold every segment is balanced.

Lemma 5.9 *Let $H_p(x, y)$ be a hinge in M with $\text{Per}(p, x, y) < 2D_\kappa$, and suppose that for every pair of points in $px \cup py$ there is a connecting segment in M . If every hinge formed by two such segments satisfies (A^κ) , then $H_p(x, y)$ satisfies (C^κ) . The analogous result for (A_κ) and (C_κ) holds if the segments px, py are balanced.*

Proof: Let $(\bar{p}, \bar{x}, \bar{y})$ be a comparison triple in $\mathbb{M}_\kappa^2$ for (p, x, y) , and let $u, v \neq p$ and $\bar{u}, \bar{v}$ be given as in (C^κ) . 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 (A^κ) by assumption, thus

$$\angle_u^\kappa(p, y) + \angle_u^\kappa(x, y) \geq \angle_u(p, y) + \angle_u(x, y) \geq \pi. \quad (1)$$

Lemma 5.3 then shows that $\angle_p^\kappa(u, y) \leq \angle_{\bar{p}}(\bar{x}, \bar{y}) = \angle_{\bar{p}}(\bar{u}, \bar{y})$, and Lemma 5.2 yields $|uy| \leq |\bar{u}\bar{y}|$. Now an analogous argument shows that $|uv| \leq |\bar{u}\bar{v}|$ if $(\bar{p}, \bar{u}, \bar{v})$ is a comparison triple for (p, u, y) and $\bar{v} \in \bar{p}\bar{y}$ is such that $|pv| = |\bar{p}\bar{v}|$. Since $|\bar{u}\bar{y}| = |uy| \leq |\bar{u}\bar{y}|$, we get that $\angle_{\bar{p}}(\bar{u}, \bar{v}) = \angle_{\bar{p}}(\bar{u}, \bar{y}) \leq \angle_{\bar{p}}(\bar{u}, \bar{y}) = \angle_{\bar{p}}(\bar{u}, \bar{v})$ and hence $|\bar{u}\bar{v}| \leq |\bar{u}\bar{y}|$ by Lemma 5.2. Thus $|uv| \leq |\bar{u}\bar{v}|$, as required.

The corresponding result for (A_κ) and (C_κ) 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. $\square$

Definition 5.10 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 $\text{Per}(p, x, y) < 2D_\kappa$ satisfies (C_κ) or (C^κ) , 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.

Lemma 5.11 *In an Alexandrov space M 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 (C_κ) . Let $(\bar{u}, \bar{v}, \bar{w})$ be a comparison triple in $\mathbb{M}_\kappa^2$ 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 $\angle_u^\kappa(q, w) \geq \angle_{\bar{u}}(\bar{q}, \bar{w}) = \angle_{\bar{u}}(\bar{v}, \bar{w})$ by Lemma 5.2. Now Lemma 5.3 shows that $\angle_q^\kappa(u, w) + \angle_q^\kappa(v, w) \leq \pi$. Letting u, v, w tend to q we get that $\angle_q(p, y) + \angle_q(x, y) \leq \pi$. $\square$

Theorem 5.12 *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 $\text{sec} \geq \kappa$ or $\text{sec} \leq \kappa$, respectively.*

Proof: Suppose first that $\sec \geq \kappa$ or $\sec \leq \kappa$. Whenever $p \in M$ and $B_r \subset TM_p$ is a normal ball, where $0 < r \leq D_\kappa$, then by Corollary 3.21 every hinge $H_p(x, y)$ in M with $x, y \in B(p, \frac{r}{2})$ satisfies (H_κ) or (H^κ) , respectively, hence also (A_κ) or (A^κ) by Lemma 5.8. Given a point $q \in M$, we can choose r so that $B_r \subset TM_p$ is normal for all $p \in B(q, \frac{r}{4})$ (compare the proof of Corollary 1.22). Now it follows from Lemma 5.9 (and since segments in a Riemannian manifold are balanced) that every hinge with vertex and endpoints in $U_q := B(q, \frac{r}{8})$ satisfies (C_κ) or (C^κ) , respectively.

Suppose now that M has curvature $\geq \kappa$ or $\leq \kappa$ in the sense of Alexandrov. Let $P \subset TM_p$ be a linear 2-plane in some tangent space, and let $\gamma_r \subset P$ be the circle with center 0 and radius $r > 0$ (sufficiently small). It follows by hinge comparison (H_κ) or (H^κ) that the length of $\exp_p(\gamma_r)$ is at most or at least

$$2\pi \operatorname{sn}_\kappa(r) = 2\pi \left(r - \frac{\kappa}{3!} r^3 + O(r^5) \right),$$

the length of a circle of radius r in $\mathbb{M}_\kappa^2$. On the other hand,

$$L(\exp_p(\gamma_r)) = 2\pi \left(r - \frac{\sec(P)}{3!} r^3 + O(r^4) \right)$$

in any Riemannian manifold (exercise). It follows that $\sec(P) \geq \kappa$ or $\sec(P) \leq \kappa$, respectively. $\square$

5.3 Toponogov's theorem

We start with two technical propositions. For both, $\kappa \in \mathbb{R}$ is a fixed constant, and M is assumed to 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.

Proposition 5.13 *Let $H_0 = H_{p_0}(x_0, y_0)$ be a hinge in M with sides of length $a_0 := |p_0 y_0|$ and $b_0 := |p_0 x_0| < \frac{1}{5}a_0$, where $a_0 + b_0 < D_\kappa$. If every hinge in M with perimeter $< \frac{4}{5} \operatorname{Per}(H_0)$ and an endpoint in $\{x_0, y_0\}$ satisfies (A_κ) , then H_0 satisfies (A_κ) as well.*

Proof: Starting from H_0 , we will inductively construct a 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 $l_n := |p_n x_n| + |p_n y_n|$ satisfy

$$a_0 + b_0 = l_0 \geq l_1 \geq l_2 \geq \dots \geq |x_0 y_0|; \quad (1)$$

furthermore, for $n \geq 1$, $|p_n x_n| = b := \frac{2}{5}a_0$ and hence

$$|p_n y_n| \geq |x_n y_n| - |p_n x_n| = |x_0 y_0| - b \geq a_0 - b_0 - b > b. \quad (2)$$

The hinge H_0 is already given. Let γ_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 $p_n x_n$ of $p_{n-1} y_{n-1}$ and an arbitrary segment $p_n y_n$. Let γ_n denote the angle of H_n , and note that the adjacent angle between $p_n y_n$ and the segment $p_n p_{n-1} \subset p_{n-1} y_{n-1}$ equals $\pi - \gamma_n$, because segments in M are balanced. Clearly (1) holds. Note also that

$$\text{Per}(p_{n-1}, p_n, y_n) \leq 2(l_{n-1} - b) \leq 2(a_0 + b_0 - b) < \frac{8}{5}a_0 \leq \frac{4}{5}\text{Per}(H_0). \quad (3)$$

Now we will construct a sequence of hinges $\bar{H}_n := H_{\bar{p}_n}(\bar{x}_n, \bar{y}_n)$ in $\mathbb{M}_\kappa^2$ such that $|\bar{p}_n \bar{x}_n| = |p_n x_n|$, $|\bar{p}_n \bar{y}_n| = |p_n y_n|$,

$$|\bar{x}_0 \bar{y}_0| \geq |\bar{x}_1 \bar{y}_1| \geq |\bar{x}_2 \bar{y}_2| \geq \dots, \quad (4)$$

and such that the angle $\bar{\gamma}_n$ of $\bar{H}_n$ is greater than or equal to γ_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 $\bar{\omega}_n := \angle_{\bar{p}_{n-1}}(\bar{p}_n, \bar{y}_n) = \angle_{\bar{p}_{n-1}}(\bar{x}_n, \bar{y}_n)$. 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. This shows (4).

Now, if $n \rightarrow \infty$, then

$$|\bar{p}_{n-1} \bar{p}_n| + |\bar{p}_{n-1} \bar{y}_n| - |\bar{p}_n \bar{y}_n| = l_{n-1} - l_n \rightarrow 0$$

by (1), consequently $\bar{\omega}_n \rightarrow \pi$ and $\bar{\gamma}_n \rightarrow \pi$ (note that $|\bar{p}_{n-1} \bar{p}_n| = |p_{n-1} y_{n-1}| - b \geq a_0 - b_0 - 2b > 0$ by (2), and $|\bar{p}_{n-1} \bar{y}_n| = b > 0$). This implies in turn that

$$l_n - |\bar{x}_n \bar{y}_n| = |\bar{p}_n \bar{x}_n| + |\bar{p}_n \bar{y}_n| - |\bar{x}_n \bar{y}_n| \rightarrow 0$$

as $n \rightarrow \infty$ (recall that $l_n \leq a_0 + b_0 < D_\kappa$). In view of (1) and (4), this gives $|\bar{x}_0 \bar{y}_0| \geq |x_0 y_0|$, so H_0 satisfies (H_κ) and hence also (A_κ) . $\square$

Proposition 5.14 *Let $H = H_p(x, y)$ be a hinge in M with $\text{Per}(H) < 2D_\kappa$. If every hinge in M with perimeter $< \frac{4}{5}\text{Per}(H)$ and an endpoint in $px \cup \{y\}$ satisfies (A_κ) , then H satisfies (A_κ) as well.*

Proof: We assume that $y \notin px$; otherwise $\angle_p^\kappa(x, y) = 0$ and (A_κ) holds trivially. For all $x' \in px$, we have $|x'y| \leq \frac{1}{2}\text{Per}(H) < D_\kappa$, we can thus choose $\delta \in (0, \frac{1}{5}d(y, px))$ such that $|x'y| + \delta < D_\kappa$ for all $x' \in px$. It follows from Proposition 5.13 that every hinge $H_{x'}(x'', y)$ with $x'x'' \subset px$ and $|x'x''| \leq \delta$ (hence $|x'x''| < \frac{1}{5}|x'y|$) satisfies (A_κ) . Subdivide px into finitely many subsegments $px_1, x_1x_2, \dots, x_{n-1}x_n, x_nx$ of length $\leq \delta$, and choose segments $x_1y, \dots, x_ny$. Put $x_0 := p$. We show by downward induction that for $i \in \{0, \dots, n\}$, the hinge $H_i = H_{x_i}(x, y)$ with $x_i x \subset px$ satisfies (A_κ) . Since $|x_n x| \leq \delta$, this holds for H_n . Suppose now that $i \geq 1$ and H_i satisfies (A_κ) . Since also $H_{x_i}(x_{i-1}, y)$ satisfies (A_κ) , and px is balanced, we have

$$\angle_{x_i}^\kappa(x_{i-1}, y) + \angle_{x_i}^\kappa(x, y) \leq \angle_{x_i}(x_{i-1}, y) + \angle_{x_i}(x, y) = \pi.$$

Furthermore, $H_{x_{i-1}}(x_i, y)$ satisfies (A_κ) , and so

$$\angle_{x_{i-1}}(x, y) = \angle_{x_{i-1}}(x_i, y) \geq \angle_{x_{i-1}}^\kappa(x_i, y) \geq \angle_{x_{i-1}}^\kappa(x, y),$$

where the last step follows from Lemma 5.3. This shows that H_{i-1} satisfies (A_κ) . $\square$

Theorem 5.15 (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 $H = H_p(x, y)$ in M with $\text{Per}(H) < 2D_\kappa$ satisfies (A_κ) , (H_κ) , and (C_κ) .*

Note that this applies, of course, to complete Riemannian manifolds with sectional curvature greater than or equal to κ . For this case, the result was established in [To1959] (see also [To1957] and [To1958]). Purely metric proofs were given later in [Pl1991], [BuGP1992], and [Pl1996].

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 (A_κ) . Suppose to the contrary that there exists a hinge H in M with $\text{Per}(H) < 2D_\kappa$ that does not satisfy (A_κ) . Then, by Proposition 5.14, there exists a hinge H_1 with $\text{Per}(H_1) < \frac{4}{5} \text{Per}(H)$ and an endpoint on the union of the sides of H such that H_1 does not satisfy (A_κ) either. Inductively, for $n = 2, 3, \dots$, there exist hinges H_n such that $\text{Per}(H_n) < \frac{4}{5} \text{Per}(H_{n-1}) < \left(\frac{4}{5}\right)^n \text{Per}(H)$, some endpoint of H_n lies on the union of the sides of H_{n-1} , and H_n does not satisfy (A_κ) . 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 (A_κ) . This gives a contradiction, as $p_n \rightarrow q$ and $\text{Per}(H_n) \rightarrow 0$. $\square$

Remark 5.16 (generalized angle and hinge comparisons) Proposition 5.14 and Theorem 5.15 can be adapted to *generalized hinges* H , where one of the sides, px , is the trace of a (locally minimizing) geodesic $c: [0, l] \rightarrow M$ of length $0 < l \leq |py| + |xy|$, and $\text{Per}(H) := l + |py| + |xy| < 2D_\kappa$. In the proof of Proposition 5.14, one can arrange that all subsegments of c of length $\leq \delta$ are minimizing. Then (A_κ) and (H_κ) hold for H if comparison triples or hinges in $\mathbb{M}_\kappa^2$ are chosen such that $|\bar{p}\bar{x}|$ or $|\hat{p}\hat{x}|$ equals l , respectively.

Theorem 5.17 *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 $\text{Per}(p, x, y) \leq 2D_\kappa$; in particular $\text{diam}(M) \leq D_\kappa$. If $\text{Per}(p, x, y) = 2D_\kappa$ and $|px|, |py| \in (0, D_\kappa)$, then there are unique segments px and py , and $\angle_p(x, y) = \pi$.*

Proof: First we show that $\text{diam}(M) \leq D_\kappa$. 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 $\text{Per}(p, q, y), \text{Per}(q, x, y) < 2D_\kappa$. Consider two segments $\bar{p}\bar{q}$ and $\bar{q}\bar{x}$ in $\mathbb{M}_\kappa^2$ such that $|\bar{p}\bar{q}| = |\bar{q}\bar{x}| = \frac{1}{2}|px|$ and $\angle_{\bar{q}}(\bar{p}, \bar{x}) = \pi$, 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 $\text{Per}(\bar{p}, \bar{y}, \bar{x}) < 2D_\kappa = \text{Per}(\bar{p}, \bar{q}, \bar{x})$ 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 (H_κ)), 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 $\text{diam}(M) \leq D_\kappa$.

Next, suppose that there exists a hinge $H_p(x, y)$ in M with $P := \text{Per}(p, x, y) > 2D_\kappa$. Since $|px|, |py|, |xy| \leq \text{diam}(M) \leq D_\kappa < \frac{1}{2}P$, there is a point $q \in px \setminus \{p, x\}$ such that $|yp| + |pq| = \frac{1}{2}P$. Let $\kappa' \in (0, \kappa)$ be such that $P = 2D_{\kappa'}$, and let $(\bar{p}, \bar{x}, \bar{y})$ be a comparison triple in $\mathbb{M}_{\kappa'}^2$, for (p, x, y) , lying on a great circle. 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 by chord comparison $(C_{\kappa''})$ for $\kappa'' < \kappa'$ and a limit argument that $|qy| \geq \frac{1}{2}P > D_\kappa$, in contradiction to $\text{diam}(M) \leq D_\kappa$. This shows that all triples in M have perimeter at most $2D_\kappa$.

Finally, suppose that $P := \text{Per}(p, x, y) = 2D_\kappa$ and $|px|, |py| \in (0, D_\kappa)$. Note that $\angle_p^\kappa(x, y) = \pi$. Hence, by angle comparison $(A_{\kappa'})$ for $\kappa' < \kappa$ and a limit argument, $\angle_p(x, y) = \pi$ for any choice of segments px, py . It then follows easily that px, py are in fact unique. $\square$

5.4 CAT(0) spaces

Next we discuss some basic results for complete geodesic metric spaces satisfying (C^0) globally. Recall first that if M is a Hadamard manifold, then every pair of points $x, y \in M$ is connected by a unique segment $xy \subset M$, and all geodesics in M are minimizing (Theorem 4.18). For two geodesics $c, \bar{c}: I \rightarrow M$ defined on the same interval, the distance function $s \mapsto d(c(s), \bar{c}(s))$ is convex (see Lemma 4.19). Furthermore, by Corollary 3.21, every hinge in M satisfies (H^0) and, hence, also (A^0) and (C^0) (Lemma 5.8 and Lemma 5.9).

Definition 5.18 For $\kappa \in \mathbb{R}$, a metric space M is called a $\text{CAT}(\kappa)$ space (or, briefly, M is $\text{CAT}(\kappa)$) if M is complete, every pair of points at distance $< D_\kappa$ in M is connected by a segment, and every hinge H with $\text{Per}(H) < 2D_\kappa$ satisfies (C^κ) (and, hence, (A^κ) and (H^κ)).

Every Hadamard manifold with sectional curvature $\leq \kappa \leq 0$ is a $\text{CAT}(\kappa)$ space. The acronym was coined in [Gro1987] (see p. 106 and p. 119 therein) and stands

for *Comparison inequality of Alexandrov–Toponogov* or *Cartan–Alexandrov–Toponogov*. In some literature (for example, [BrH1999]), completeness is not part of the definition. However, the comparison properties are preserved under metric completion (see Corollary II.3.11 in [BrH1999]), and completeness is required for most results below, so we prefer to include it from the beginning.

It is not difficult to see that if M is a $\text{CAT}(\kappa)$ space, then every pair of points at distance $< D_\kappa$ in M is connected by a *unique* segment, all balls of radius $< D_\kappa/2$ are convex, and balls of radius $< D_\kappa$ are contractible (see Proposition II.1.4 in [BrH1999]). A metric space is $\text{CAT}(\kappa)$ if and only if it is $\text{CAT}(\kappa')$ for every $\kappa' > \kappa$ (see Theorem II.1.12 in [BrH1999]). Although this will not be pursued in these notes, we mention that $\text{CAT}(1)$ and $\text{CAT}(0)$ spaces occur notably as (completed) spaces of directions and tangent cones, respectively, of spaces of curvature $\leq \kappa$ in the sense of Alexandrov (see Theorem II.3.19 in [BrH1999] and the references therein).

We now focus on $\text{CAT}(0)$ spaces. First we note that the convexity of the distance function mentioned above (Lemma 4.19) is retained in this more general context.

Remark 5.19 (Busemann convexity) If M is a $\text{CAT}(0)$ space and $c, \bar{c}: I \rightarrow M$ are two geodesics defined on the same interval, then the distance function $s \mapsto |c(s) \bar{c}(s)|$ is convex. To see this, assume that $I = [0, 1]$, and let $\gamma: [0, 1] \rightarrow M$ be the geodesic from $c(0)$ to $\bar{c}(1)$. The $\text{CAT}(0)$ property yields $|c(t) \gamma(t)| \leq t |c(1) \bar{c}(1)|$ as well as $|\gamma(t) \bar{c}(t)| \leq (1 - t) |c(0) \bar{c}(0)|$, and so

$$|c(t) \bar{c}(t)| \leq (1 - t) |c(0) \bar{c}(0)| + t |c(1) \bar{c}(1)|.$$

A complete geodesic metric space M with this convexity property is called *Busemann convex* or a *Busemann space* (see [Bus1948] and [Jos1997], [Pap2005]).

Every Banach space with a strictly convex norm is a Busemann space; however, the $\text{CAT}(0)$ property holds only in the case of a Hilbert space (see Proposition 5.22 below).

Proposition 4.20 (flat strip) also holds for $\text{CAT}(0)$ or Busemann spaces (with essentially the same proof): if $c, \bar{c}: \mathbb{R} \rightarrow M$ are two unit speed geodesics such that $s \mapsto |c(s) \bar{c}(s)|$ is bounded, then the convex hull of $c(\mathbb{R}) \cup \bar{c}(\mathbb{R})$ is isometric to a strip $\mathbb{R} \times [0, D] \subset \mathbb{R}^2$ with respect to the Euclidean or some strictly convex norm, respectively. The discussion of isometries in Section 4.5, 4.21–4.25, also remains valid for $\text{CAT}(0)$ and Busemann spaces. We refer to Chapters II.7 and II.8 in [BrH1999] for a comprehensive exposition in the $\text{CAT}(0)$ case.

Remark 5.20 (Local-to-global) Local-to-global results, akin to Theorem 5.15 (Toponogov), hold for simply connected spaces of non-positive curvature: a complete, simply connected, and locally Busemann convex length space is (globally) Busemann convex, and a complete, simply connected length space of curvature ≤ 0

in the sense of Alexandrov (equivalently, satisfying the CAT(0) condition locally) is CAT(0) (see [AleB1990] and Chapter II.4 in [BrH1999]). In either case, it follows that the underlying space is contractible, in analogy with Theorem 4.18 (Hadamard–Cartan).

The CAT(0) condition is often used in the form of one of the following inequalities (compare [BruT1972, Sect. 3.2]).

Proposition 5.21 *For a complete geodesic metric space X , the following are equivalent:*

- (1) X is a CAT(0) space.
- (2) If $p \in X$ and $c: [0, 1] \rightarrow X$ is a minimizing geodesic from x to y , then

$$|p c(t)|^2 \leq (1-t) |px|^2 + t |py|^2 - t(1-t) |xy|^2$$

for all $t \in [0, 1]$.

- (3) If $p, x, y \in X$ and z is a midpoint between x and y , then

$$|pz|^2 \leq \frac{1}{2} (|px|^2 + |py|^2) - \frac{1}{4} |xy|^2.$$

Furthermore, if X is a complete metric space, and if for all pairs of points $x, y \in X$ there exists a point $z \in X$ such that the last inequality holds for all $p \in X$, then z is the unique midpoint between x and y , and X is CAT(0).

Proof: Suppose that X is a CAT(0) space. To show that (2) holds, choose vectors $u, v \in \mathbb{R}^2$ such that $\|u\| = |xp|$, $\|v\| = |xy|$, and $\|u - v\| = |py|$. Chord comparison for the hinge $H_x(p, y)$ gives $|p c(t)| \leq \|u - tv\|$, and since

$$\|u - tv\|^2 - t \|u - v\|^2 = (1-t) \|u\|^2 + (t^2 - t) \|v\|^2,$$

(2) follows. Conversely, suppose that (2) holds. Then $|p c(t)| \leq \|u - tv\|$. Likewise, if z is a point on some segment xp with $|xz| = s|xp|$, $s \in [0, 1]$, and $w \in \mathbb{R}^2$ is chosen such that $\|w\| = |x c(t)|$ and $\|u - w\| = |p c(t)|$, then $|z c(t)| \leq \|su - w\|$. Since $\|u - w\| \leq \|u - tv\|$, it follows from the law of cosines that also $\|su - w\| \leq \|su - tv\|$. Thus $|z c(t)| \leq \|su - tv\|$. This shows that chord comparison (C^0) holds in X .

Clearly (2) implies the *midpoint inequality* (3). For the reverse implication, we first show that if the inequality in (2) holds for two parameter values $0 \leq s < u \leq 1$, then it also holds for the mean $t := (s + u)/2$. Indeed, by (3) and the inequalities for s and u ,

$$\begin{aligned} |p c(t)|^2 &\leq \frac{1}{2} (|p c(s)|^2 + |p c(u)|^2) - \frac{1}{4} |c(s) c(u)|^2 \\ &\leq (1-t) |px|^2 + t |py|^2 - T |xy|^2, \end{aligned}$$

where $T = (s(1-s) + u(1-u))/2 + (u-s)^2/4 = t(1-t)$. It now follows easily that (2) holds for all dyadic rationals in $[0, 1]$ and hence for all $t \in [0, 1]$.

For the last assertion of the proposition, note that the inequality in (3) applied for $p = x$ and $p = y$ shows that z is at distance at most $|xy|/2$ from both x and y and is thus a midpoint. Since X is complete, it is then not difficult to show that X is a geodesic metric space, and the result follows from the implication (3) $\Rightarrow$ (1). $\square$

Proposition 5.22 *If $(X, \|\cdot\|)$ is a Banach space such that X is CAT(0) with respect to the induced metric, then X is in fact a Hilbert space, and the inner product satisfies $\langle x, y \rangle = (\|x + y\|^2 - \|x - y\|^2)/4$.*

Proof: For $x, y \in X$, the midpoint inequality yields

$$\|x + y\|^2 = 4 \left\| \frac{x + y}{2} \right\|^2 \leq 2(\|x\|^2 + \|y\|^2) - \|x - y\|^2$$

as well as

$$2\|x\|^2 = 2 \left\| \frac{(x + y) + (x - y)}{2} \right\|^2 \leq \|x + y\|^2 + \|x - y\|^2 - 2\|y\|^2,$$

so that X satisfies the *parallelogram law*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Hence, by the well-known characterization of pre-Hilbert spaces due to Jordan–von Neumann [JoN1935], the norm of X is induced by an inner product, which in turn satisfies $\langle x, y \rangle = (\|x + y\|^2 - \|x - y\|^2)/4$. $\square$

The parallelogram law continues to hold as an inequality for arbitrary quadruples in a CAT(0) space. This follows from work of Reshetnyak (see [Re1968], [KoS1993], [Jos1997]).

Proposition 5.23 (quadrilateral inequality) *Let X be a CAT(0) space. For every quadruple of points $p, q, x, y \in X$,*

$$\begin{aligned} |pq|^2 + |xy|^2 &\leq 2|px||qy| + |py|^2 + |qx|^2 \\ &\leq |px|^2 + |qy|^2 + |py|^2 + |qx|^2. \end{aligned}$$

In fact, this inequality, in the form with the six squares, characterizes CAT(0) spaces among all complete geodesic spaces, see [BeN2007] and [Sat2009].

Proof: Let $c: [0, 1] \rightarrow X$ be the geodesic from x to y . For all $t \in (0, 1)$,

$$|pq|^2 \leq (|p c(t)| + |q c(t)|)^2 \leq \frac{1}{t} |p c(t)|^2 + \frac{1}{1-t} |q c(t)|^2.$$

Together with the second assertion of Proposition 5.21, this gives the inequality

$$\begin{aligned} |pq|^2 &\leq \frac{1-t}{t} |px|^2 + |py|^2 - (1-t) |xy|^2 \\ &\quad + |qx|^2 + \frac{t}{1-t} |qy|^2 - t |xy|^2. \end{aligned}$$

Assuming that $p \neq x$ and $q \neq y$, we now take $t := |px|/(|px| + |qy|)$. Then $(1-t)/t = |qy|/|px|$, and the result follows. $\square$

A subset C of a CAT(0) space X is *convex* if for every pair of points $x, y \in C$ the segment xy is contained in C .

Proposition 5.24 (metric projection) *Suppose that X is a CAT(0) space, and $C \subset X$ is a non-empty closed convex set. Then for all $p \in X$ there exists a unique point $\pi(p) \in C$ with $d(p, \pi(p)) = d(p, C)$, and the resulting map $\pi: X \rightarrow C$ is 1-Lipschitz.*

Proof: Let $p \in X$, and put $a := d(p, C)$. Whenever $x, y \in C$, and $c: [0, 1] \rightarrow X$ is the geodesic from x to y , then $c(t) \in C$ for all $t \in (0, 1)$, hence

$$a^2 \leq |p c(t)|^2 \leq (1-t) |px|^2 + t |py|^2 - t(1-t) |xy|^2. \quad (1)$$

For $t = \frac{1}{2}$, this gives $|xy|^2 \leq 2(|px|^2 + |py|^2 - 2a^2)$, thus any sequence (x_i) in C with $|px_i| \rightarrow a$ is Cauchy, and the limit is the desired unique projection point $\pi(p) \in C$.

Let now $q \in X$ be another point, and put $b := d(q, C)$. For $x := \pi(p)$ and $y := \pi(q)$, (1) yields $a^2 + (1-t) |xy|^2 \leq |py|^2$ for all $t \in (0, 1)$ (note that $|px| = a$), hence $a^2 + |xy|^2 \leq |py|^2$. Likewise, $b^2 + |xy|^2 \leq |qx|^2$. Using Proposition 5.23, we conclude that

$$a^2 + b^2 + 2 |xy|^2 \leq |py|^2 + |qx|^2 \leq a^2 + b^2 + |pq|^2 + |xy|^2,$$

thus $|xy| \leq |pq|$. This shows that π is 1-Lipschitz. $\square$

The next result shows that every non-empty bounded subset of a CAT(0) space has a unique *circumcenter*.

Proposition 5.25 (circumcenter) *Let X be a CAT(0) space. If $A \subset X$ is a non-empty bounded set, and*

$$r := \inf\{s \geq 0 : \text{there exists } x \in X \text{ such that } A \subset \bar{B}(x, s)\},$$

then there is unique point $x \in X$ such that $A \subset \bar{B}(x, r)$.

Proof: Whenever $x, y \in X$ and $s, t \geq 0$ are such that $A \subset \bar{B}(x, s) \cap \bar{B}(y, t)$, and z is the midpoint between x and y , then the definition of r and the midpoint inequality give

$$4r^2 \leq \sup\{4|pz|^2 : p \in A\} \leq 2(s^2 + t^2) - |xy|^2,$$

thus $|xy|^2 \leq 2(s^2 + t^2 - 2r^2)$. It follows that any sequence (x_i) in X with $A \subset \bar{B}(x_i, s_i)$ and $s_i \rightarrow r$ is Cauchy, and the limit point is the unique circumcenter of A . $\square$

The following consequence was observed in [BruT1972].

Corollary 5.26 *Suppose that X is a CAT(0) space and $\Gamma \subset \text{Isom}(X)$ is a group of isometries with bounded orbits $\Gamma(p) = \{\gamma(p) : \gamma \in \Gamma\}$. Then the fixed point set*

$$\text{Fix}(\Gamma) = \{x \in X : \gamma(x) = x \text{ for all } \gamma \in \Gamma\}$$

is non-empty and convex.

Proof: Take a point $p \in X$, and let x be the unique circumcenter of the bounded set $A := \Gamma(p)$. For every $\gamma \in \Gamma$, the point $\gamma(x)$ is a circumcenter of $\gamma(A) = A$, thus $\gamma(x) = x$ by uniqueness. This shows that $x \in \text{Fix}(\Gamma)$.

For two points $x, y \in \text{Fix}(\Gamma)$, every $\gamma \in \Gamma$ fixes the endpoints of the (unique) segment xy and must therefore fix xy pointwise. Thus $\text{Fix}(\Gamma)$ is convex. $\square$

The analogue of Corollary 5.26 for Busemann convex spaces is not true in general. In fact, there exists a bounded Busemann space admitting an isometry without fixed points; see [Bas2018].

We conclude this section with a standard application of Corollary 5.26 in geometric group theory.

Corollary 5.27 *Suppose that X is a CAT(0) space and $G \subset \text{Isom}(X)$ is a group of isometries such that for some subset $B \subset X$, the set $\{g \in G : g(B) \cap B \neq \emptyset\}$ is finite and $\bigcup_{g \in G} g(B) = X$. Then G contains only finitely many conjugacy classes of finite subgroups.*

This applies in particular to properly discontinuous and cocompact isometric actions.

Proof: Put $F := \{g \in G : g(B) \cap B \neq \emptyset\}$. By Corollary 5.26, every finite subgroup $\Gamma \subset G$ fixes a point $x \in X$. If $g \in G$ is such that $x \in g(B)$, then $g^{-1}\Gamma g$ fixes the point $g^{-1}(x) \in B$, and so $g^{-1}\Gamma g \subset F$. Hence every finite subgroup of G has a conjugate contained in the finite set F . $\square$

CAT(0) spaces will be discussed further in Section 8.1.

Chapter 6

Volume comparison and applications

6.1 Volume comparison theorems

For $r > 0$ we denote by $V_{m,\kappa}(r)$ the volume of a ball of radius r in $\mathbb{M}_\kappa^m$. Explicitly,

$$V_{m,\kappa}(r) = \omega_{m-1} \int_0^{\inf\{r, D_\kappa\}} \text{sn}_\kappa^{m-1}(t) dt,$$

where ω_{m-1} is the $(m-1)$ -dimensional volume of the unit sphere in $\mathbb{R}^m$.

Theorem 6.1 (Günther 1960) *Let M be a Riemannian m -manifold with $\text{sec} \leq \kappa$ 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*

$$\text{vol}(B(p, r)) \geq V_{m,\kappa}(r).$$

Proof: This follows easily from Corollary 3.20. See [Gü1960]. $\square$

For a map $F: M \rightarrow M'$ 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 ,

$$JF(q) = \frac{\sqrt{\det(\langle dF_q(b_i), dF_q(b_j) \rangle)}}{\sqrt{\det(\langle b_i, b_j \rangle)}}.$$

We define the positive function $J_{m,\kappa}: [0, D_\kappa) \rightarrow \mathbb{R}$ by

$$J_{m,\kappa}(r) := \frac{\text{sn}_\kappa^{m-1}(r)}{r^{m-1}}$$

for $r \in (0, D_\kappa)$, and $J_{m,\kappa}(0) := 1$. This is the Jacobian of the exponential map $\exp_p: T(\mathbb{M}_\kappa^m)_p \rightarrow \mathbb{M}_\kappa^m$, for any $p \in \mathbb{M}_\kappa^m$, at any point $v \in T(\mathbb{M}_\kappa^m)_p$ with $|v| = r$.

The following result, as well as Theorem 6.6 below, appeared in [BisC1964].

Proposition 6.2 (Bishop 1964) *Let M be a Riemannian m -manifold. Suppose that $0 < l \leq \infty$ and*

$$\varrho: [0, l) \rightarrow M, \quad \varrho(r) = \exp_p(ru),$$

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 $\text{Ric}(\varrho'(r), \varrho'(r)) \geq \kappa(m-1)$ for all $r \in [0, l)$, for some constant $\kappa \in \mathbb{R}$. Then $l \leq D_\kappa$, and the function

$$f: [0, l) \rightarrow \mathbb{R}, \quad f(r) = \frac{J \exp_p(ru)}{J_{m, \kappa}(r)},$$

satisfies $1 = f(0) \geq f(r) \geq f(s) > 0$ whenever $0 \leq r \leq s < l$. 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_κ . We want to show that $f'(r) \leq 0$. Let $Y_1, \dots, Y_{m-1}$ be Jacobi fields along ϱ 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 3.11). They are of the form

$$Y_i(t) = d(\exp_p)_{tu}(tv_i)$$

for a basis $(v_1, \dots, v_{m-1})$ of $u^\perp \subset TM_p$ ($v_i = Y_i'(0)$, see the proof of Lemma 3.12). Note that $J \exp_p(tu) = e(t)/\bar{e}(t)$ for $t \in (0, l)$, where

$$e(t) := \sqrt{\det(\langle Y_i(t), Y_j(t) \rangle)} \quad \text{and} \quad \bar{e}(t) := \sqrt{\det(\langle tv_i, tv_j \rangle)} = t^{m-1} \bar{e}(1)$$

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

$$e'(r) = \frac{1}{2} \frac{d}{dt} \Big|_{t=r} \det(\langle Y_i(t), Y_j(t) \rangle) = \sum_{i=1}^{m-1} \langle Y_i(r), Y_i'(r) \rangle.$$

Since Y_i is a Jacobi field with $Y_i(0) = 0$, the term $\langle Y_i(r), Y_i'(r) \rangle$ agrees with $I'(Y_i, Y_i)$, the index form of $\varrho|_{[0, r]}$. Let now E_i denote the parallel vector field along ϱ with $E_i(r) = Y_i(r)$, and put $\lambda := \text{sn}_\kappa(r)$ and $h := \frac{1}{\lambda} \text{sn}_\kappa$. Proposition 3.17 (first index lemma) then yields the inequality

$$\begin{aligned} \langle Y_i(r), Y_i'(r) \rangle &= I^r(Y_i, Y_i) \\ &\leq I^r(hE_i, hE_i) = \int_0^r (h')^2 - h^2 R(E_i, \varrho', E_i, \varrho') dt. \end{aligned}$$

Using the assumption on the Ricci curvature we deduce that

$$e'(r) \leq (m-1) \int_0^r (h')^2 - \kappa h^2 dt.$$

By the choice of hE_i , equality holds for the model space $\mathbb{M}_\kappa^m$, where $e(t)$ becomes $h^{m-1}(t)$. Thus $e'(r) \leq (h^{m-1})'(r)$. Since

$$f(t) = \frac{J \exp_p(tu)}{J_{m,\kappa}(t)} = \frac{e(t)}{\bar{e}(t)} \frac{t^{m-1}}{\text{sn}_\kappa^{m-1}(t)} = \frac{1}{\bar{e}(1) \lambda^{m-1}} \frac{e(t)}{h^{m-1}(t)}$$

and $e(r) = 1 = h^{m-1}(r)$, it follows that $f'(r) \leq 0$.

To show that $l \leq D_\kappa$, suppose to the contrary that $D_\kappa < l$. Then the above argument still shows that f is non-increasing on $[0, D_\kappa)$. However, $J_{m,\kappa}(r) \rightarrow 0$ for $r \rightarrow D_\kappa$, whereas $J \exp_p(ru) > 0$ for $r \in [0, l)$, because ϱ 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 $I'(Y_i, Y_i) = I'(hE_i, hE_i)$ in the above argument, for $i = 1, \dots, m-1$. Then Proposition 3.17 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 $\text{sn}_\kappa E$ is a Jacobi field. Thus, by the Jacobi equation,

$$\langle \text{sn}_\kappa E, R(\text{sn}_\kappa E, \varrho') \varrho' \rangle = \langle \text{sn}_\kappa E, -\text{sn}_\kappa'' E \rangle = \kappa \text{sn}_\kappa^2.$$

It follows that $\sec(\text{span}\{\varrho'(t), E(t)\}) = \kappa$ for all $t \in [0, r]$. $\square$

Definition 6.3 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

$$\text{Cut}_p := \{t_u u : u \in TM_p, |u| = 1, t_u < \infty\} \subset TM_p$$

defines the *tangent cut locus* of p , and the set $\text{Cut}(p) := \exp_p(\text{Cut}_p)$ of *cut points* of p is the *cut locus* of p . The *injectivity radius* of M at p is defined as

$$\text{inj}_p := d(p, \text{Cut}(p)) = \inf\{t_u : u \in TM_p, |u| = 1\},$$

and $\text{inj}(M) := \inf\{\text{inj}_p : p \in M\}$ is the *injectivity radius* of M .

Lemma 6.4 Suppose that M is complete. Let $c_u : \mathbb{R} \rightarrow M$, $c_u(t) := \exp_p(tu)$, 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 a unit vector $v \in TM_p$ distinct from u such that $c_v(t) = c_u(t)$.

Conversely, if (1) or (2) holds for some $t \in (0, \infty)$, then $t_u \leq t$.

Proof: Choose a decreasing sequence $t_i \rightarrow t = t_u$, and let $\gamma_i: [0, l_i] \rightarrow M$ be a minimizing unit speed geodesic from p to $c_u(t_i)$. We may assume that the vectors $v_i := \gamma_i'(0)$ converge to a unit vector v . Then $c_v|_{[0, l]}$ is a minimizing geodesic from p to $c_u(t)$ for $l := \lim_{i \rightarrow \infty} l_i$. Since also $c_u|_{[0, t]}$ is minimizing, $l = t$. In case $v \neq u$, (2) holds. Suppose now that $v = u$. Since

$$\exp_p(t_i u) = c_u(t_i) = \gamma_i(l_i) = \exp_p(l_i v_i)$$

and $t_i u \rightarrow t u$ as well as $l_i v_i \rightarrow t u$, it follows that $\exp_p$ must be singular at $t u$, for otherwise $t_i u = l_i v_i$ and $d(p, \exp_p(t_i u)) = d(p, \gamma_i(l_i)) = l_i = t_i$ for large i , in contradiction to the definition of the cut value. Hence, by Lemma 3.12, (1) holds.

Conversely, if (1) holds for some $t > 0$, then $d(p, c_u(t')) < t'$ for all $t' > t$ according to Theorem 3.16, so $t_u \leq t$. Suppose now that (2) holds for some $t > 0$, and let $t' > t$. Since $v \neq u$, the geodesics $c_v|_{[0, t]}$ and $c_u|_{[t, t']}$ form an angle less than π at $c_v(t) = c_u(t)$, and therefore $d(p, c_u(t'))$ is smaller than the sum of their lengths, $t + (t' - t) = t'$. Thus $t_u \leq t$. $\square$

Proposition 6.5 *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 $U_p := \{tu : 0 \leq t < t_u\}$ is open, $\partial U_p = \text{Cut}_p$, and $\exp_p|_{U_p}$ is a diffeomorphism from U_p onto the open set $M \setminus \text{Cut}(p)$.*
- (3) *$\text{Cut}(p)$ is a set of measure zero in M .*

Proof: We prove (1). Suppose that $u_i \rightarrow u$, and put $t_i := t_{u_i}$. Whenever $0 < r < \limsup_{i \rightarrow \infty} t_i$, there are infinitely many i such that $t_i > r$ and hence $d(p, c_{u_i}(r)) = r$; then $d(p, c_u(r)) = r$ and thus $r \leq t_u$. This shows that $\limsup_{i \rightarrow \infty} t_i \leq t_u$. We may assume that $s := \liminf_{i \rightarrow \infty} t_i < \infty$. Then it follows from Lemma 6.4 and Lemma 3.12 that there is a subsequence $t_{i_k} \rightarrow s$ such that either $\exp_p$ is singular at $t_{i_k} u_{i_k}$ for all k , or

$$\exp_p(t_{i_k} v_k) = \exp_p(t_{i_k} u_{i_k})$$

for some convergent sequence $v_k \rightarrow v$ of unit vectors $v_k \neq u_{i_k}$. In the first case, $\exp_p$ is singular at su because the set of singular points is closed, so $t_u \leq s$ by the same two lemmas. Now consider the second case. If $v \neq u$, then $t_u \leq s$ by Lemma 6.4 since $c_v(s) = c_u(s)$. If $v = u$, then $t_{i_k} v_k \neq t_{i_k} u_{i_k}$ both converge to su , which is therefore a singular point of $\exp_p$, so that again $t_u \leq s$. Hence, $t_i \rightarrow t_u$. $\square$

Theorem 6.6 (Bishop 1964) *Let M be a Riemannian m -manifold with $\text{Ric} \geq (m - 1)\kappa g$ 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*

$$\text{vol}(B(p, r)) \leq V_{m, \kappa}(r).$$

If M is complete and $\kappa > 0$, then $\text{vol}(M) \leq \text{vol}(\mathbb{M}_\kappa^m)$, and equality holds only if M is isometric to $\mathbb{M}_\kappa^m$.

Proof: If $r > 0$ is such that $B_r \subset TM_p$ is a normal ball, then it follows from Proposition 6.2 that $r \leq D_\kappa$ and

$$\text{vol}(B(p, r)) = \int_{B_r} J \exp_p(v) dv \leq \int_{B_r} J_{m,\kappa}(|v|) dv = V_{m,\kappa}(r).$$

Suppose now that M is complete. If $0 < r \leq D_\kappa$, 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 $\text{vol}(M) \leq \text{vol}(\mathbb{M}_\kappa^m)$, because then $\text{diam}(M) \leq r = \text{diam}(\mathbb{M}_\kappa^m)$ and $U_p \subset B_r$. In particular, the inequality $\text{vol}(B(p, r)) \leq V_{m,\kappa}(r)$ also holds for $r > D_\kappa$.

To prove the last assertion, suppose that M is complete, $\kappa > 0$, and $\text{vol}(M) = \text{vol}(\mathbb{M}_\kappa^m)$. 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 κ . As M and $\mathbb{M}_\kappa^m$ have equal volume, they must be isometric (recall Theorem 4.13). $\square$

The following result was observed in [Gro1981a].

Theorem 6.7 (Bishop–Gromov 1981) *Let M be a complete Riemannian m -manifold with $\text{Ric} \geq (m-1)\kappa g$ for some constant $\kappa \in \mathbb{R}$, and let $p \in M$. Then the function*

$$h: (0, \infty) \rightarrow \mathbb{R}, \quad h(r) = \frac{\text{vol}(B(p, r))}{V_{m,\kappa}(r)},$$

satisfies $h(r) \geq h(s)$ whenever $0 < r \leq s < \infty$, and $h(r) \rightarrow 1$ as $r \rightarrow 0$.

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 ω_{m-1} . Let b be the corresponding function for $\mathbb{M}_\kappa^m$, 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

$$\text{vol}(B(p, r)) = \int_0^r a(t) dt \quad \text{and} \quad V_{m,\kappa}(r) = \int_0^r b(t) dt.$$

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 $h(r) \rightarrow 1$ as $r \rightarrow 0+$. $\square$

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].

Theorem 6.8 (Cheng 1975) *Let M be a complete Riemannian m -manifold with $\text{Ric} \geq (m-1)\kappa g$ for some constant $\kappa > 0$, and with (maximal possible) diameter $\text{diam}(M) = D_\kappa$. Then M is isometric to $\mathbb{M}_\kappa^m$.*

Proof: For $p \in M$ and $r > 0$, set $h_p(r) := \text{vol}(B(p, r))/V_{m, \kappa}(r)$. Now fix $p, \bar{p} \in M$ such that $d(p, \bar{p}) = \text{diam}(M) = D_\kappa$. 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

$$\text{vol}(M) \geq \text{vol}(B(p, r)) + \text{vol}(B(\bar{p}, \bar{r})) = h_p(r) V_{\kappa, m}(r) + h_{\bar{p}}(\bar{r}) V_{\kappa, m}(\bar{r}).$$

By Theorem 6.7,

$$h_p(r) \geq h_p(D_\kappa) = \frac{\text{vol}(M)}{\text{vol}(\mathbb{M}_\kappa^m)}, \quad h_{\bar{p}}(\bar{r}) \geq h_{\bar{p}}(D_\kappa) = \frac{\text{vol}(M)}{\text{vol}(\mathbb{M}_\kappa^m)}.$$

Since $V_{\kappa, m}(r) + V_{\kappa, m}(\bar{r}) = \text{vol}(\mathbb{M}_\kappa^m)$, it follows that equality holds in all of these estimates, in particular $h_p(r) = \text{vol}(M)/\text{vol}(\mathbb{M}_\kappa^m)$, independently of r . Since $\lim_{r \rightarrow 0} h_p(r) = 1$, this shows that $\text{vol}(M) = \text{vol}(\mathbb{M}_\kappa^m)$. Hence, by Theorem 6.6, M is isometric to $\mathbb{M}_\kappa^m$. $\square$

6.2 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 Γ , that is, every element $\gamma \in \Gamma$ can be written as a

finite product $\gamma = a_0 \cdots a_n$, where $a_0 = e$ and $a_1, \dots, a_n \in A \cup \{a^{-1} : a \in A\}$. The smallest number $n \geq 0$ for which such a representation of γ exists is called the *word length* of γ 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, Γ 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 Γ , 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 Γ 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.

Definition 6.9 The *growth function* of Γ 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 Γ has *polynomial growth of degree at most $k \geq 0$* if there exists a constant $c \geq 1$ such that

$$w_A(r) \leq cr^k \quad \text{for all } r,$$

and Γ 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 Γ 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 $d_{A'} \leq L d_A$ as above, then $w_A(r) \leq w_{A'}(Lr)$ for all r . We remark further that $w_A(r+s) \leq w_A(r) w_A(s)$, so the upper exponential bound $w_A(r) \leq w_A(1)^r$ is always satisfied.

For example, the free abelian group Γ on k generators $a_1, \dots, a_k$ has polynomial growth of degree k . For $k \geq 2$, the free group Γ on k generators $a_1, \dots, a_k$ has exponential growth.

Lemma 6.10 Let (M, g) be a closed Riemannian manifold with universal covering $(\tilde{M}, \tilde{g})$, and let $\Gamma \subset \text{Isom}(\tilde{M}, \tilde{g})$ 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

$$A := \{a \in \Gamma : d(K, aK) < \delta\}$$

is finite and generates Γ ; in fact, if $\gamma \in \Gamma$ and $n \geq 1$ is an integer such that $d(K, \gamma K) < n\delta$, then $|\gamma|_A \leq n$.

In particular, the fundamental group $\pi_1(M) \simeq \Gamma$ is finitely generated.

Proof: Since the action of Γ on $\tilde{M}$ is properly discontinuous, A is finite. Now let $\gamma \in \Gamma$, and let $n \geq 1$ be an 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

$$\gamma = (\gamma_0^{-1} \gamma_1)(\gamma_1^{-1} \gamma_2) \cdots (\gamma_{n-1}^{-1} \gamma_n) = a_1 a_2 \cdots a_n.$$

Furthermore, $d(K, a_i K) = d(\gamma_{i-1} K, \gamma_i K) \leq d(p_{i-1}, p_i) < \delta$, thus $a_i \in A$ and $|\gamma|_A \leq n$. $\square$

Theorem 6.11 (Milnor 1968) *Let M be a closed Riemannian manifold with $\sec < 0$. Then $\pi_1(M)$ has exponential growth.*

Proof: We continue with the notation of Lemma 6.10. Let $p \in K$ and $n \geq 1$. If $\gamma \in \Gamma$ and $\gamma K \cap B(p, n\delta) \neq \emptyset$, then $d(K, \gamma K) < n\delta$ and hence $|\gamma|_A \leq n$ by the lemma. Thus $B(p, n\delta)$ is covered by $w_A(n)$ (or less) translates of K , all of volume $\text{vol}(K)$, and so

$$\text{vol}(B(p, n\delta)) \leq w_A(n) \text{vol}(K).$$

Furthermore, since $\sec < 0$ and M is compact, there is a constant $\kappa < 0$ such that $\sec \leq \kappa$. By Theorem 4.18 (Hadamard–Cartan), $B_{n\delta} \subset T\tilde{M}_p$ is a normal ball, so $\text{vol}(B(p, n\delta)) \geq V_{m,\kappa}(n\delta)$ by Theorem 6.1 (Günther). Since $V_{m,\kappa}$ is exponential, $\pi_1(M) \simeq \Gamma$ has exponential growth. $\square$

Theorem 6.11 is no longer true in general if the bound on the sectional curvature is replaced by the assumption $\text{Ric} < 0$: by a result of Lohkamp [Lo1992], [Lo1994], for every $m \geq 3$ there exist constants $a_m > b_m > 0$ such that every m -dimensional manifold M admits a complete Riemannian metric with $-a_m(m-1)g < \text{Ric} < -b_m(m-1)g$.

Theorem 6.12 (Milnor 1968) *Let M be a complete Riemannian m -manifold with $\text{Ric} \geq 0$. Then every finitely generated subgroup of $\pi_1(M)$ has polynomial growth of degree at most m .*

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

$$d(p, \gamma(p)) \leq \mu |\gamma|_A \quad \text{for all } \gamma \in \Gamma'.$$

Let $\epsilon > 0$ be such that $d(p, \gamma(p)) \geq 2\epsilon$ for all $\gamma \in \Gamma' \setminus \{e\}$. For every integer $r \geq 0$, the ball $B(p, \mu r + \epsilon)$ contains all balls $B(\gamma(p), \epsilon)$ with $|\gamma|_A \leq r$, and the latter are pairwise disjoint by the choice of ϵ . As there are $w_A(r)$ such ϵ -balls, all of the same volume, we get that

$$w_A(r) \operatorname{vol}(B(p, \epsilon)) \leq \operatorname{vol}(B(p, \mu r + \epsilon)) \leq V_{m,0}(\mu r + \epsilon)$$

by Theorem 6.6 (Bishop). It follows that Γ' has polynomial growth of degree at most m . $\square$

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 4.26).

For another application, consider the Heisenberg group $\tilde{M} = H$ of real upper triangular 3×3 -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 $\operatorname{Ric} \geq 0$.

Milnor conjectured that under the assumptions of Theorem 6.12, the fundamental group $\pi_1(M)$ is finitely generated. This holds in dimension 3 [Liu2013], [Pa2020]. Wilking [Wi2000] proved that if Γ is a group such that for some m , every finitely generated subgroup has polynomial growth of degree at most m (as $\Gamma = \pi_1(M)$ in Theorem 6.12), then Γ is finitely generated if and only if every abelian subgroup is finitely generated. However, very recently, Bruè, Naber and Semola [BruNS2025] proved Milnor's conjecture to be wrong by constructing a 7-dimensional counterexample.

For complete manifolds of non-negative *sectional* curvature, the following holds [Gro1978].

Theorem 6.13 (Gromov 1978) *For every complete m -dimensional Riemannian manifold with $\sec \geq 0$, 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 Γ' . First, pick $a_1 \in \Gamma' \setminus \{e\}$ so that $d_1 := d(p, a_1(p)) \leq d(p, a(p))$ for all $a \in \Gamma' \setminus \{e\}$, and let Γ_1 denote the subgroup of Γ' 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 Γ_{j-1} they generate is still smaller than Γ' . Then pick $a_j \in \Gamma' \setminus \Gamma_{j-1}$ such that

$$d_j := d(p, a_j(p)) \leq d(p, a(p)) \quad \text{for all } a \in \Gamma' \setminus \Gamma_{j-1},$$

and let Γ_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

$$d_{ij} := d(a_i(p), a_j(p)) = d(p, (a_i^{-1}a_j)(p)) \geq d_j;$$

furthermore $d_j \geq d_i$ 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 $\angle_p(a_i(p), a_j(p))$ is greater than or equal to the corresponding angle in a Euclidean triangle with sides of lengths $d_i \leq d_j \leq d_{ij}$, 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 $\geq \pi/3$, or the maximal number of pairwise disjoint open balls of radius $\frac{1}{2}$ contained in $B_{3/2} \subset TM_p$, which is less than or equal to 3^m . $\square$

6.3 Gromov–Hausdorff convergence

For subsets A, B of a metric space $X = (X, d)$ we denote by

$$N_\delta(A) = \{x \in X : d(x, A) \leq \delta\}$$

the closed δ -neighborhood of A and by

$$d_H(A, B) = \inf\{\delta \geq 0 : A \subset N_\delta(B), B \subset N_\delta(A)\} \in [0, \infty]$$

the *Hausdorff distance* of A and B ; d_H defines a metric on the set $\mathcal{C}$ 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 ϵ . 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 [B11916].

Theorem 6.14 *Suppose that X is a metric space and $\mathcal{C}$ is the set of non-empty, closed and bounded subsets of X , endowed with the Hausdorff metric d_H .*

(1) *If X is complete, then $\mathcal{C}$ is complete.*

(2) *If X is compact, then $\mathcal{C}$ is compact.*

Proof: For the proof (1), let $(B_i)_{i \in \mathbb{N}}$ be a Cauchy sequence in $\mathcal{C}$. 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 $d_H(B_i, B_j) < \epsilon/2$ 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

$$d(x, B_i) \leq d(x, B_j) + d_H(B_i, B_j) < \epsilon \quad \text{for all } i \geq i_0.$$

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 $i = i_1 < i_2 < \dots$ such that $d_H(B_m, B_n) < \epsilon/2^k$ 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

$$d(x, y) = \lim_{k \rightarrow \infty} d(x, x_k) \leq \sum_{k=1}^{\infty} d(x_k, x_{k+1}) < \epsilon,$$

and y belongs to $\overline{\bigcup_{i \geq k} B_{i_i}}$ 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 $\mathcal{C}$ is precompact. Let $\epsilon > 0$. There exists a finite set $A \subset X$ such that $N_\epsilon(A) = X$. We show that every $B \in \mathcal{C}$ is at Hausdorff distance at most ϵ 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 have $d_H(B, A_B) \leq \epsilon$. As there are only finitely many distinct subsets of A , it follows that $\mathcal{C}$ is precompact. $\square$

The next definition originates from [Gro1981b].

Definition 6.15 The *Gromov–Hausdorff distance* $d_{\text{GH}}(X, Y) \in [0, \infty]$ 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 $d_H^Z(X', Y') \leq \delta$.

For example, if both X and Y have diameter less than or equal to $D \geq 0$, then $d_{\text{GH}}(X, Y) \leq \frac{1}{2}D$: choose (Z, d^Z) as 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 $(x, y) \in X' \times Y'$ (notice that d^Z satisfies the triangle inequality).

Proposition 6.16 (1) d_{GH} satisfies the triangle inequality, that is, for all metric spaces X, Y, Z ,

$$d_{\text{GH}}(X, Z) \leq d_{\text{GH}}(X, Y) + d_{\text{GH}}(Y, Z).$$

(2) d_{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]. $\square$

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 $(\mathcal{C}, d_H)$ might suggest). Let X be the geodesic tree with one central vertex and edges of length $1 - \frac{1}{2}, 1 - \frac{1}{3}, 1 - \frac{1}{4}, \dots$ attached to it, and define Y similarly, but with an additional segment of length 1. Then $d_{GH}(X, Y) = 0$ 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_∞ 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 $x_{k(1)}, x_{k(2)}, \dots$ 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 $\sup_{\alpha \in A} \text{diam}(X_\alpha) < \infty$ and for all $\epsilon > 0$ there exists an integer $n(\epsilon)$ such that each X_α can be covered by $n(\epsilon)$ open balls of radius ϵ .

Theorem 6.17 (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_α 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 $i \in \mathbb{N} := \{1, 2, \dots\}$, put $\epsilon_i := 2^{-i}$ and let $n_i \in \mathbb{N}$ be such that each X_α can be covered by n_i open balls of radius ϵ_i . Fix a partition $\mathbb{N} = \bigcup_{i=1}^\infty N_i$ and a map $\pi: \mathbb{N} \setminus N_1 \rightarrow \mathbb{N}$ such that the set N_i has $n_1 n_2 \dots n_i$ 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_α , construct a sequence $(x_k^\alpha)_{k \in \mathbb{N}}$ in the following way. For $i = 1$, choose x_k^α for $k \in N_1$ so that the n_1 balls $B(x_k^\alpha, \epsilon_1)$ cover X_α . For $i \geq 1$, if the $n_1 n_2 \dots n_i$ points x_k^α with $k \in N_i$ are already selected, choose x_l^α for $l \in N_{i+1}$ so that for every $k \in N_i$, the n_{i+1} balls $B(x_l^\alpha, \epsilon_{i+1})$ with $l \in \pi^{-1}\{k\}$ cover $B(x_k^\alpha, \epsilon_i)$ and are contained in $B(x_k^\alpha, 2\epsilon_i)$. This inductive process yields for each $\alpha \in A$ a dense sequence $(x_k^\alpha)_{k \in \mathbb{N}}$ in X_α . Let $f_\alpha: X_\alpha \rightarrow l_\infty$ 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 $l \in \pi^{-1}\{k\}$,

$$|d(x, x_k^\alpha) - d(x, x_l^\alpha)| \leq d(x_k^\alpha, x_l^\alpha) < 2\epsilon_i.$$

In particular, 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 $0 \leq s_k \leq \sup_\alpha \text{diam}(X_\alpha) < \infty$ for all $k \in \mathbb{N}$ and

$$|s_k - s_l| \leq 2\epsilon_i \quad \text{whenever } i \in \mathbb{N} \text{ and } \pi(l) = k \in N_i.$$

It follows that if $i \in \mathbb{N}$ and $l \in N_j$ for some $j > i$, and $k := \pi^{j-i}(l) \in N_i$, then $|s_k - s_l| \leq \sum_{j=i}^\infty 2\epsilon_j = 4\epsilon_i$. Now it is not difficult to see that $Z \subset l_\infty$ is precompact. Hence, as a closed subset of a Banach space, Z is compact.

(2) follows directly from (1) and the second part of Theorem 6.14. $\square$

Theorem 6.18 (Gromov) *For $m \geq 2$, $\kappa \in \mathbb{R}$, and $D > 0$, the set $\mathcal{M}$ of isometry classes of closed Riemannian m -manifolds M with $\text{Ric} \geq \kappa(m-1)g$ and $\text{diam}(M) \leq D$ 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 $\text{vol}(B(z, \epsilon/2)) \leq \text{vol}(B(p, \epsilon/2))$ for all $p \in A$. Then $\text{vol}(M) \geq |A| \text{vol}(B(z, \epsilon/2))$, and therefore

$$|A| \leq \frac{\text{vol}(M)}{\text{vol}(B(z, \epsilon/2))} = \frac{\text{vol}(B(z, D))}{\text{vol}(B(z, \epsilon/2))} \leq \frac{V_{m,\kappa}(D)}{V_{m,\kappa}(\epsilon/2)} =: n(\epsilon)$$

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 ϵ . The second assertion follows from Theorem 6.17. $\square$

Remark 6.19 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 1.26).
- (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.)

6.4 Diffeomorphism finiteness

In this section we explain, following [Gro1981c], how Theorem 6.18 can be used for the proof of a finiteness theorem due to Cheeger [Ch1970] (see also [Pet1984]). We start with a result from [Kl1959].

Proposition 6.20 (Klingenberg) *Let M be a closed Riemannian manifold.*

- (1) *Let $p \in M$, and let q be a point in $\text{Cut}(p)$ with minimal distance to p ,*

$$l := d(p, q) = d(p, \text{Cut}(p)) = \text{inj}_p.$$

If q is not conjugate to p along any minimizing geodesic from p to q , then there exist exactly two distinct minimizing geodesics $\sigma, \tau: [0, l] \rightarrow M$ from p to q , and $\sigma'(l) = -\tau'(l)$.

- (2) *If $\text{sec} \leq \kappa$ and $\text{inj}(M) < D_\kappa$ for some $\kappa > 0$, then there exists a closed geodesic in M of length $2 \text{inj}(M)$.*

Proof: We prove (1). If q is not conjugate to p along any minimizing geodesic, then by Lemma 6.4 there exist two distinct minimizing geodesics $\sigma, \tau: [0, l] \rightarrow M$ from p to q , and it suffices to show that necessarily $\sigma'(l) = -\tau'(l)$. Suppose to the contrary that $\sigma'(l) \neq -\tau'(l)$. Then $\langle w, \sigma'(l) \rangle, \langle w, \tau'(l) \rangle < 0$ for some vector $w \in TM_q$. Since $\exp_p$ is regular at both $l\sigma'(0)$ and $l\tau'(0)$ (as q is not conjugate to p along σ or τ), there exists an open neighborhood V of q in M such that $\exp_p$ maps some disjoint neighborhoods of $l\sigma'(0)$ and $l\tau'(0)$ diffeomorphically onto V . Choose a curve $c: (-\epsilon, \epsilon) \rightarrow V$ with $c(0) = q$ and $c'(0) = w$. Then there exists a variation of $\sigma = \sigma_0$ through geodesics $\sigma_s: [0, l] \rightarrow M$ from p to $c(s)$. By the first variation formula, $\frac{d}{ds}\big|_{s=0} L(\sigma_s) = \langle w, \sigma'(l) \rangle < 0$, thus $L(\sigma_s) < L(\sigma) = l$ for small $s > 0$. Analogously, there exists a variation of $\tau = \tau_0$ through geodesics $\tau_s: [0, l] \rightarrow M$ from p to $c(s)$ with $L(\tau_s) < l$ for small $s > 0$. This contradicts the injectivity of $\exp_p$ on U_p (see Proposition 6.5).

For (2), note that since M is compact, there exist $p \in M$ and $q \in \text{Cut}(p)$ such that $\text{inj}_p = \text{inj}(M)$ and $l := d(p, q) = d(p, \text{Cut}(p)) = \text{inj}_p$. Since $\text{sec} \leq \kappa$, Theorem 3.19 (Rauch) shows that there are no conjugate points along geodesics of length $l = \text{inj}(M) < D_\kappa$. Hence, by (1), there exist two minimizing geodesics $\sigma, \tau: [0, l] \rightarrow M$ from p to q such that $\sigma'(l) = -\tau'(l)$. Furthermore, by Lemma 6.4, the cut value of $-\sigma'(l) \in TM_q$ is less than or equal to $l = \text{inj}(M)$, hence equal to l , and so $p \in \text{Cut}(q)$ and $d(p, q) = \text{inj}_q$. Now it follows again from (1) that also $-\sigma'(0) = \tau'(0)$. $\square$

Proposition 6.21 (Cheeger) *Given $m \geq 2$, $\kappa \in \mathbb{R}$, and $D, V > 0$, there exists a constant $l > 0$ such that if M is a closed Riemannian m -manifold with $\text{sec} \geq \kappa$, $\text{diam}(M) \leq D$, and $\text{vol}(M) > V$, then every non-constant closed geodesic in M has length greater than l .*

Proof: In case $\kappa > 0$ we suppose without loss of generality that $D \leq D_\kappa$. Let $\sigma: \mathbb{R} \rightarrow M$ be a closed unit speed geodesic with period $b > 0$. Put $p := \sigma(0)$ and $u := \sigma'(0)$. For $r \in (0, D)$ and $\alpha \in (0, \pi/2)$, consider the open subset

$$W_{r,\alpha} := B_D \setminus \{v : |v| \geq r, \angle(u, v) \in [0, \alpha] \cup [\pi - \alpha, \pi]\}$$

of TM_p . Fix a linear isometry $H: TM_p \rightarrow T(\mathbb{M}_\kappa^m)_q$ for some $q \in \mathbb{M}_\kappa^m$. Since $V < \text{vol}(M) \leq V_{m,\kappa}(D)$ by Theorem 6.6, we can choose r and α such that

$$\text{vol}(\exp_q(H(W_{r,\alpha}))) = V.$$

Let $U_p \subset TM_p$ be as in Proposition 6.5. By Proposition 6.2,

$$V < \text{vol}(M) = \text{vol}(\exp_p(U_p)) \leq \text{vol}(\exp_q(H(U_p))).$$

Hence, $U_p \not\subset W_{r,\alpha}$. Since $U_p \subset B_D$, it follows that there exists a minimizing unit speed geodesic $\varrho: [0, r] \rightarrow M$ with $\varrho(0) = p$ such that $\angle(u, \varrho'(0))$ is $\leq \alpha$ or $\geq \pi - \alpha$. By reversing σ if necessary, we arrange that $\angle(u, \varrho'(0)) \leq \alpha$. Recall that $\alpha < \pi/2$. Now it follows by generalized hinge comparison (see Remark 5.16) that there exists a constant $l > 0$, depending only on κ, r, α (where r, α depend only on m, κ, 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 $b > l$. □

Theorem 6.22 (Gromov) *Given $m \geq 2$ and $\kappa, \varrho > 0$, there exists a constant $\delta > 0$ such that if M, M' are two closed Riemannian m -manifolds with*

$$|\text{sec}_M|, |\text{sec}_{M'}| \leq \kappa, \quad \text{inj}(M), \text{inj}(M') \geq \varrho, \quad \text{and} \quad d_{\text{GH}}(M, M') < \delta,$$

then M and M' are diffeomorphic.

Proof: See Section 8.D in [Gro1981c] and the references therein. □

Theorem 6.23 (Cheeger 1970) *Given $m \geq 2$ and $\kappa, D, V > 0$, there exist only finitely many diffeomorphism classes of closed Riemannian m -manifolds M with $|\text{sec}_M| \leq \kappa$, $\text{diam}(M) \leq D$, and $\text{vol}(M) \geq V$.*

Proof: Proposition 6.21 provides a positive lower bound for the length of non-constant closed geodesics, and by the second part of Proposition 6.20 there exists a constant $\varrho > 0$ such that $\text{inj}(M) \geq \varrho$ for every such M . Now the result follows from Theorem 6.18 and Theorem 6.22. □

Chapter 7

Open manifolds of non-negative curvature

We will now discuss the global shape of *open* (that is, complete and non-compact) Riemannian manifolds of non-negative sectional curvature. Some general references are [GrM1969], [ChG1972], [Sh1974], Section 8 in [ChE1975], and [Pe1994].

We start with a brief general discussion of Busemann functions.

7.1 Busemann functions

Let M be a complete Riemannian manifold. A *ray* ϱ in M is a geodesic $\varrho: [0, \infty) \rightarrow M$, 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 ϱ with $\varrho(0) = p$. To see this, choose a sequence of minimizing geodesics $\varrho_i: [0, d(p, q_i)] \rightarrow M$ from p to q_i , where $d(p, q_i) \rightarrow \infty$; then some subsequence converges uniformly on compact sets to a ray emanating from p .

Now let a ray ϱ in M be given. For $x \in M$ and $t \geq s \geq 0$,

$$d(x, \varrho(t)) - t \leq d(x, \varrho(s)) + d(\varrho(s), \varrho(t)) - t = d(x, \varrho(s)) - s,$$

furthermore the function $t \mapsto d(x, \varrho(t)) - t$ is bounded from below by $-d(x, \varrho(0))$. Therefore the limit

$$b_\varrho(x) := \lim_{t \rightarrow \infty} d(x, \varrho(t)) - t$$

exists. This defines the *Busemann function* $b_\varrho: M \rightarrow \mathbb{R}$ of ϱ . Note that $b_\varrho(\varrho(s)) = -s$ for all $s \geq 0$. If $x, y \in M$, then $|d(x, \varrho(t)) - d(y, \varrho(t))| \leq d(x, y)$ for all $t \geq 0$ and thus

$$|b_\varrho(x) - b_\varrho(y)| \leq d(x, y);$$

that is, b_ϱ 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*. Note that

$$B(\varrho, r) = \bigcup_{t \in [0, \infty)} B(\varrho(t), r + t);$$

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 $t \geq 0$.

Lemma 7.1 *Let M be a complete Riemannian manifold. Given a ray ϱ and a point q in M , there exists a ray σ such that $\sigma(0) = q$ and*

$$b_\varrho(\sigma(s)) = b_\varrho(q) - s$$

for all $s \geq 0$. The Busemann function b_σ of any such ray σ satisfies

$$b_\sigma(x) \geq b_\varrho(x) - b_\varrho(q)$$

for every $x \in M$.

Note that, since b_ϱ is 1-Lipschitz, the inequality $b_\varrho(c(s)) \geq b_\varrho(q) - s$ holds for any unit speed curve $c: [0, \infty) \rightarrow M$ with $c(0) = q$. The ray σ is thus a curve of steepest descent for the function b_ϱ .

Proof: For every $t \geq 0$, let $\sigma_t: [0, d(q, \varrho(t))] \rightarrow M$ be a minimizing geodesic from q to $\varrho(t)$. Then, for a fixed $s \geq 0$ and for all sufficiently large t , since b_ϱ is 1-Lipschitz,

$$b_\varrho(q) - s \leq b_\varrho(\sigma_t(s)) \leq b_\varrho(\varrho(t)) + d(\sigma_t(s), \varrho(t)) = -t + d(q, \varrho(t)) - s.$$

As $t \rightarrow \infty$, the last term tends to $b_\varrho(q) - s$, hence $b_\varrho(\sigma_t(s)) \rightarrow b_\varrho(q) - s$ as well. For some sequence $t_i \rightarrow \infty$ the σ_{t_i} converge, uniformly on compact sets, to the desired ray σ .

Let $x \in M$. 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 σ satisfies $b_\sigma(x) \geq b_\varrho(x) - b_\varrho(q)$, as claimed. $\square$

Let further M denote a complete Riemannian manifold. A set $C \subset M$ is called *totally convex* if every geodesic $c: [0, 1] \rightarrow M$ (minimizing or not) with $c(0), c(1) \in C$ satisfies $c(\lambda) \in C$ for all $\lambda \in [0, 1]$.

A function $f: C \rightarrow \mathbb{R}$ on a totally convex set $C \subset M$ is called *concave* if for every geodesic $c: [0, 1] \rightarrow C$ and every $\lambda \in [0, 1]$,

$$f(c(\lambda)) \geq (1 - \lambda)f(c(0)) + \lambda f(c(1)).$$

Note that then, for every $r \in \mathbb{R}$, the superlevel set $\{f \geq r\}$ is totally convex. To check that f is concave, since geodesics are locally minimizing, it suffices to verify the above inequality for every *minimizing* geodesic $c: [0, 1] \rightarrow C$.

Lemma 7.2 *Let M be a complete Riemannian manifold of non-negative sectional curvature, and let ϱ be a ray in M . Then the Busemann function $b_\varrho: M \rightarrow \mathbb{R}$ 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 $c: [0, 1] \rightarrow M$ be a minimizing geodesic from x to $y \neq x$. Let $\lambda \in (0, 1)$, and put $q := c(\lambda)$. By Lemma 7.1 there exists a ray σ with $\sigma(0) = q$ whose Busemann function satisfies

$$b_\sigma(z) \geq b_\varrho(z) - b_\varrho(q)$$

for all $z \in M$. Let γ 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 γ with $\bar{q}\bar{y}$. By Theorem 5.15 (hinge comparison (H_0)), $d(x, \sigma(s)) \leq d(\bar{x}, \bar{\sigma}(s))$ for all $s \geq 0$, and 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 is zero since $b_{\bar{\sigma}}$ is affine and $b_{\bar{\sigma}}((1-\lambda)\bar{x} + \lambda\bar{y}) = b_{\bar{\sigma}}(\bar{q}) = 0$. $\square$

7.2 The soul construction

We now turn to the so-called soul construction in complete non-compact Riemannian manifolds of non-negative curvature.

Proposition 7.3 *Let M be an open 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 $r < s \leq t$, the set C_r contains the closed $(s-r)$ -neighborhood of C_s . Hence the family (C_r) exhausts M , that is, $\bigcup_r C_r = M$. The curvature assumption will only be used through Lemma 7.2.

Proof: Let R denote the set of all rays starting at the given point p . For every $r \in \mathbb{R}$, define the closed set

$$C_r := \bigcap_{\varrho \in R} (M \setminus B(\varrho, r)) = \{x \in M : b_\varrho(x) \geq r \text{ for all } \varrho \in R\}.$$

By Lemma 7.2, C_r is totally convex. For every ray $\varrho \in R$, $\varrho(0) = p$ is a boundary point of $M \setminus B(\varrho, 0)$, so $p \in \partial C_0$. Evidently $C_s \subset C_r$ 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 $d(p, q_i) \rightarrow \infty$ for $i \rightarrow \infty$. Since also $p \in C_r$ and C_r is totally convex, for every q_i there is a minimizing geodesic $\varrho_i: [0, d(p, q_i)] \rightarrow M$ 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 7.1 there exists a ray σ 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 $l \in \{0, \dots, m\}$ 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 .

Proposition 7.4 *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 $c: [0, 1] \rightarrow M$ 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 (recall Proposition 2.21). 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 of L restricted to $\{v \in TL^\perp : |v| < \epsilon\}$ is a diffeomorphism onto some neighborhood V of p in M and furthermore $B_\epsilon \subset TM_q$ is a normal ball for every $q \in V$. Note that

$$L \subset L(C) \cap V \subset C \cap V.$$

Suppose that there exists a point $q \in (C \cap V) \setminus L$. Then there is an $x \in L$ and a segment qx of length $< \epsilon$ perpendicular to L . For $y \in L \cap B(q, \epsilon)$ and $\lambda \in (0, 1)$, the map sending (y, λ) to $\exp_q(\lambda(\exp_q|_{B_\epsilon})^{-1}(y))$ is smooth and immersive along $\{x\} \times (0, 1)$, furthermore the image lies in C because $q \in C$ and $L \subset C$. Thus there is a non-empty $(l+1)$ -dimensional submanifold of M contained in C , in 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 $c: [a, b] \rightarrow M$ 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 $t \in (r, r + \epsilon)$, $t \leq b$. A similar argument as above, with H and $c(t)$ in place of L and q , shows that every $c(s)$ with $s \in (r, t)$ belongs to an l -dimensional submanifold of M contained in C . This proves the claim. Since $c(a) \in L(C)$, we conclude that $c([a, b]) \subset L(C)$. $\square$

For C and $L(C)$ as in Proposition 7.4, we now define the *border*

$$bC := C \setminus L(C).$$

Note that bC equals the topological boundary ∂C of C in M if and only if $\dim(C) = m = \dim(M)$. The following results supplements Proposition 7.4.

Proposition 7.5 *Let $p \in bC$. Then there is an l -dimensional linear subspace of TM_p , denoted by TC_p , such that $\exp_p^{-1}(C) \subset TC_p$ and $\exp_p^{-1}(L(C))$ is open in TC_p , furthermore there is a unit vector $n_p \in TC_p$ such that $\langle n_p, v \rangle > 0$ for all $v \in \exp_p^{-1}(L(C))$.*

The ‘inner normal’ vector n_p is uniquely determined for a dense set of points in bC , as the proof will show, but is non-unique in general.

Proof: Let $p \in bC$, and let $\tau: [0, 1] \rightarrow M$ be any geodesic with $\tau([0, 1)) \subset L(C)$ and $\tau(1) = p$. Since $L(C)$ is totally geodesic, M -parallel transport along τ preserves the family of tangent spaces $TL(C)_{\tau(s)}$ for $s \in [0, 1)$ and, hence, takes each member of this family to the same subspace of TM_p . Call this subspace TC_p .

Let $v \in TM_p$ with $\exp_p(v) \in C$. Note that $c_v([0, 1]) \subset C$. To show that $v \in TC_p$, choose $\epsilon > 0$ and $q := c_v(t)$, $t \in (0, 1]$, such that $B_\epsilon \subset TM_q$ is a normal ball and $p \in B(q, \epsilon)$. For every $s \in [0, 1]$ with $\tau(s) \in B(q, \epsilon)$, let $\gamma_s: [0, t] \rightarrow B(q, \epsilon)$ be the geodesic from $\tau(s)$ to q . If $s < 1$, then $\gamma_s([0, t)) \subset L(C)$ and $\gamma'_s(0) \in TL(C)_{\tau(s)}$, thus $v = c'_v(0) = \gamma'_1(0) \in TC_p$ by continuity.

Now let $v \in TC_p$ with $\exp_p(v) \in L(C)$. For a normal ball $B_\delta \subset TM_p$, the set $U := (\exp_p|_{B_\delta})^{-1}(L(C) \cap B(p, \delta))$ is an l -dimensional submanifold, hence open

in TC_p . Since $c_v((0, 1]) \subset L(C)$, there is a $\lambda \in (0, 1)$ such that $\lambda v \in U$ and the segment $[\lambda, 1]v$ is disjoint from the closed set $\exp_p^{-1}(bC)$. Hence, if $v' \in TC_p$ is sufficiently close to v , then $\lambda v' \in U$ and $[\lambda, 1]v' \cap \exp_p^{-1}(bC) = \emptyset$. Then $c_{v'}((0, \lambda]) \subset L(C)$, and since $L(C)$ is totally geodesic, $c_{v'}((0, 1]) \subset L(C)$; in particular $\exp_p(v') \in L(C)$. This shows that $\exp_p^{-1}(L(C))$ is open in TC_p .

For the last assertion, suppose first that $p \in bC$ is a ‘projection point’: there exists a point $q \in L(C)$ such that $d(q, p) = r := d(q, bC)$. Let $\sigma: [0, r] \rightarrow C$ be a unit speed geodesic from q to p , and put $n_p := -\sigma'(r)$. Assume, for the moment, that $v \in TC_p$ and $\langle n_p, v \rangle > 0$. Let V be the parallel vector field along σ with $V(r) = v$; note that $V(t) \in TL(C)_{\sigma(t)}$ for all $t \in [0, r]$. Consider the variation $(s, t) \mapsto \sigma_s(t) := \exp_{\sigma(t)}(s \frac{t}{r} V(t))$ of σ in M . Since $\langle n_p, v \rangle > 0$, there is an $\epsilon > 0$ such that the curves σ_s with $s \in (0, \epsilon)$ have length $< r$, thus $\sigma_s(t) \notin bC$ for all $(s, t) \in (0, \epsilon) \times [0, r]$. It follows that $[0, \epsilon) \ni s \mapsto \sigma_s(t)$ is a geodesic in $L(C)$ for every $t \in [0, r]$ and, hence, $\exp_p(sv) = \sigma_s(r) \in L(C)$ for all $s \in (0, \epsilon)$ by continuity. But then $\exp_p(-v) \notin L(C)$ because $L(C)$ is totally convex. This shows that $\langle n_p, v \rangle \geq 0$ for all $v \in \exp_p^{-1}(L(C))$, and since this set is open in TC_p , the inequality is in fact strict. Finally, for a general point $p \in bC$, pick a sequence $q_i \rightarrow p$ of points in $L(C)$. There is a corresponding sequence $p_i \rightarrow p$ of projection points in bC , and we may assume that the respective vectors $n_{p_i} \in TC_{p_i}$ converge to a unit vector $n_p \in TC_p$ (note that TC_p depends continuously on p). For any $v \in \exp_p^{-1}(L(C))$, a similar continuity argument as in the second paragraph of the proof then shows that $\langle n_p, v \rangle > 0$. $\square$

Proposition 7.6 *Let M be a complete Riemannian manifold of non-negative sectional curvature. Let C be a compact totally convex subset of M with $bC \neq \emptyset$. Then the function $f := d(\cdot, bC): C \rightarrow \mathbb{R}$ is concave, thus for every $r \in \mathbb{R}$ the superlevel set $\{f \geq r\}$ 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 $c: [0, 1] \rightarrow L(C)$ be a minimizing geodesic from x_0 to $x_1 \neq x_0$, let $\lambda \in (0, 1)$, and put $q := c(\lambda)$. Let $p \in bC$ 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

$$\gamma := \angle_p(x_0, x_1) \geq \bar{\gamma} := \angle_{\bar{p}}(\bar{x}_0, \bar{x}_1)$$

(angle comparison (A₀)) and $f(q) = d(p, q) \geq d(\bar{p}, \bar{q})$ (chord comparison (C₀)) for the hinge formed by c and the segment x_0p). 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 evidently

$$f(q) = d(p, q) = (1 - \lambda) d(p, x_0) + \lambda d(p, x_1) \geq (1 - \lambda) f(x_0) + \lambda f(x_1).$$

Suppose now that $\bar{\gamma} > 0$. For $i = 0, 1$, let $\bar{y}_i(\phi)$ be the point on the half-line $\{\bar{x}_i + s(\bar{p} - \bar{q}) : s \geq 0\}$ with $\angle_{\bar{p}}(\bar{x}_i, \bar{y}_i(\phi)) = \phi \geq 0$. Let $u_i \in TC_p$ be the initial vector of px_i with $|u_i| = d(p, x_i)$. Since $\angle(u_0, u_1) = \gamma \in (0, \pi)$, 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 $v_i(\phi) \in P_i$ the point with $|v_i(\phi)| = d(\bar{p}, \bar{y}_i(\phi))$ and $\angle(u_i, v_i(\phi)) = \phi$. 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 $z_i \in bC$. Then $\omega_0 + \gamma + \omega_1 \leq \pi$, for otherwise there would exist ϕ_0, ϕ_1 such that $\phi_0 + \gamma + \phi_1 = \pi$ and $\exp_p(v_i(\phi_i)) \in L(C)$, which would imply $p \in L(C)$ as above. Let $\bar{z}_i := \bar{y}_i(\omega_i)$. As $\omega_0 + \bar{\gamma} + \omega_1 \leq \pi$, it follows readily that

$$f(q) = d(p, q) \geq d(\bar{p}, \bar{q}) \geq (1 - \lambda) d(\bar{x}_0, \bar{z}_0) + \lambda d(\bar{x}_1, \bar{z}_1).$$

By the generalized hinge version of Toponogov's theorem (see Remark 5.16; note that we do not know whether $[0, 1] \ni t \mapsto \exp_p(tv_i(\omega_i))$ is minimizing) we get that $d(\bar{x}_i, \bar{z}_i) \geq d(x_i, z_i) \geq f(x_i)$. This gives the result. $\square$

We are now ready for the proof of the soul theorem.

Theorem 7.7 (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 Σ 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 construction of Σ results from Proposition 7.3 and Proposition 7.6 and will depend only on the choice of the point $p \in M$ in the first of these results. Any manifold Σ obtained in this way will be called a *soul* of M . 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 7.3 gives a descending family $(C_r)_{r \in (-\infty, t]}$ of non-empty compact totally convex subsets of M , where $C_s \subset C_r$ is the subset of points at distance $\geq s - r$ from ∂C_r ($r < s \leq t$), and C_t has empty interior, thus $\dim(C_t) < m$. Put $t_1 := t$ and $l_1 := \dim(C_{t_1})$. If $bC_{t_1} = \emptyset$, then $\Sigma := L(C_{t_1}) = C_{t_1}$ is already the desired totally convex submanifold.

If $bC_{t_1} \neq \emptyset$, then Proposition 7.6 gives a further family of non-empty compact totally convex sets

$$C_r := \{q \in C_{t_1} : d(q, bC_{t_1}) \geq r - t_1\}, \quad r \in (t_1, t_2],$$

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 $bC_{t_2} = \emptyset$, then we put $\Sigma := L(C_{t_2}) = C_{t_2}$.

If $bC_{t_2} \neq \emptyset$, 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 $t_1 < \dots < t_k$ and $m > l_1 > \dots > l_k \geq 0$, we have that $\dim(C_r) = l_i$ for $r \in [t_i, t_{i+1})$, $\dim(C_{t_k}) = l_k$, $bC_r \neq \emptyset$ for $r < t_k$, and $bC_{t_k} = \emptyset$. 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 Σ is compact, there exists an $\epsilon > 0$ such that the restriction of the normal exponential map of Σ to $\{u \in T\Sigma^\perp : |u| < 3\epsilon\}$ is a diffeomorphism onto the open 3ϵ -neighborhood $U_{3\epsilon}$ of Σ 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 $p \in M \setminus \Sigma$. A vector $w \in TM_p$ will be called *inward* if $\langle w, c'(0) \rangle > 0$ for every geodesic $c: [0, l] \rightarrow M$ from p to Σ of length $l := d(p, \Sigma)$. There is an $r < t_k$ such that $p \in bC_r$, and $\Sigma \subset L(C_r)$, thus by Proposition 7.5 there exists an inward unit vector $n_p \in T(C_r)_p$. Let N_p be any local C^∞ vector field with $N_p(p) = n_p$. An indirect argument, using that Σ is closed, shows that in some neighborhood V_p of p all vectors of N_p are inward. Let Y denote the (inward) gradient field of $-d(\cdot, \Sigma)$ on $U_{3\epsilon} \setminus \Sigma$. Assuming that $V_p \subset B(p, \epsilon)$, we use a partition of unity subordinate to $\{U_{3\epsilon} \setminus \Sigma\} \cup \{V_p : p \in M \setminus U_{3\epsilon}\}$ to combine these local fields to a C^∞ vector field X on $M \setminus \Sigma$ that agrees with Y on $U_{2\epsilon} \setminus \Sigma$. Every X_p is inward and hence non-zero. By a variation argument, the maximal integral curve $\sigma_p: (\alpha_p, \omega_p) \rightarrow M \setminus \Sigma$ of $\frac{1}{|X|}X$ with $\sigma_p(0) = p$ satisfies $d(\sigma_p(t), \Sigma) < d(\sigma_p(s), \Sigma)$ whenever $t > s$, and $\alpha_p = -\infty$. If $p \in U_{2\epsilon}$, then σ_p follows the shortest segment from p to Σ , and $\omega_p = d(p, \Sigma)$. Now let $p \in M \setminus U_{2\epsilon}$. If $\sigma_p([0, \omega_p)) \cap U_{2\epsilon} = \emptyset$, then $\omega_p = \infty$, and a subsequence of $(\sigma_p(n))_{n=1,2,\dots}$ would converge to a point q with $d(\sigma_q(0), \Sigma) = d(\sigma_q(1), \Sigma)$, a contradiction. This shows that σ_p eventually reaches $U_{2\epsilon}$, and again $\omega_p < \infty$.

Let now $\psi: [0, \infty) \rightarrow [0, 2\epsilon)$ be a smooth function with $\psi' > 0$ such that $\psi(t) = t$ for $t \in [0, \epsilon]$ and $\lim_{t \rightarrow \infty} \psi(t) = 2\epsilon$. Define $F: M \rightarrow U_{2\epsilon}$ such that $F(p) = p$ for $p \in \Sigma$ and $F(p) = \sigma_p(\omega_p - \psi(\omega_p))$ for $p \in M \setminus \Sigma$, so that $F(p)$ is at distance $\psi(\omega_p)$ from the soul. Note that $F(p) = \sigma_p(0) = p$ 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 Σ 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 $c: [0, 1] \rightarrow \Sigma$ be a minimizing geodesic from x_0 to x_1 as in the proof of Proposition 7.6. Recall that Σ is of the form $\{q \in C_{t_{k-1}} : d(q, bC_{t_{k-1}}) \geq t_k - t_{k-1}\}$; in the case $k = 1$, this holds for an arbitrary $t_0 < t_1$ by Proposition 7.3. Furthermore Σ has empty interior relative to $L(C_{t_{k-1}})$. It follows that $d(c(\lambda), bC_{t_{k-1}}) = t_k - t_{k-1}$ for all $\lambda \in [0, 1]$. This yields $\gamma = \bar{\gamma}$ in the proof of Proposition 7.6, and one can then conclude that $\sec(P) = 0$ for the plane $P \in TC_p$ (exercise). $\square$

7.3 The Sharafutdinov retraction

Our next goal is to show that for any soul Σ of M there exists a distance non-increasing retraction of M onto Σ , a result due to Sharafutdinov [Sh1977] (see also [Yi1988]).

For this purpose, we fix numbers $t_1 < \dots < t_k$ and a family $(C_r)_{r \in (-\infty, t_k]}$ with $\Sigma = C_{t_k}$ as in the proof of Theorem 7.7. To unify the notation, we choose an arbitrary sequence $t_0 > t_{-1} > t_{-2} > \dots \rightarrow -\infty$ with $t_0 < t_1$. Now, whenever $i \leq k$ and $t_{i-1} \leq r < s \leq t_i$, then

$$C_s = \{q \in C_r : d(q, bC_r) \geq s - r\}.$$

(In case $i \leq 1$, this holds by Proposition 7.3 as $bC_r = \partial C_r$ for $r < t_1$. If $i > 1$, then $C_s = \{q \in C_{t_{i-1}} : d(q, bC_{t_{i-1}}) \geq t_{i-1} - s\}$ for all $s \in (t_{i-1}, t_i]$ by construction, and the relation follows from the triangle inequality.) To show the existence of a 1-Lipschitz retraction $\phi: M \rightarrow \Sigma$, it suffices to construct a 1-Lipschitz retraction $\phi_i: C_{t_{i-1}} \rightarrow C_{t_i}$ for every $i \leq k$. Indeed, every point $p \in M$ is contained in some $C_{t_{i-1}}$ and one can then define $\phi(p) := (\phi_k \circ \phi_{k-1} \circ \dots \circ \phi_i)(p)$. We need two auxiliary results, in both of which the index $i \leq k$ is fixed.

Lemma 7.8 *If $t_{i-1} \leq r < s < t_i$, then*

$$d(p, C_s) \leq \frac{\text{diam}(C_s)}{t_i - s}(s - r)$$

for all $p \in C_r$.

Proof: Let $p \in C_r \setminus C_s$. Choose a point $q \in C_{t_i}$ and a minimizing geodesic $c: [0, 1] \rightarrow C_r$ from p to q . The function $h := d(c(\cdot), bC_r)$ satisfies $h(0) < s - r < t_i - r = h(1)$, thus there is a $\lambda \in (0, 1)$ such that $h(\lambda) = s - r$, and $x := c(\lambda) \in bC_s$. Since h is concave (Proposition 7.6) and $h(0) \geq 0$,

$$s - r = h(\lambda) \geq \lambda h(1) = \lambda(t_i - r),$$

hence $(1 - \lambda)(t_i - r) \geq t_i - s$ and

$$\frac{d(p, x)}{d(q, x)} = \frac{\lambda}{1 - \lambda} \leq \frac{s - r}{t_i - s}.$$

Since $q, x \in C_s$, this gives the result. $\square$

Proposition 7.9 *For every $\epsilon > 0$ there exist constants $\delta, a > 0$ such that the following holds whenever $t_{i-1} \leq r < s \leq t_i - \epsilon$ and $s - r \leq \delta$: for every $p \in C_r$ there is a unique nearest point $\psi(p)$ in C_s , and the retraction $\psi: C_r \rightarrow C_s$ so defined is c^{-1} -Lipschitz for $c := \cos(a(r - s))$.*

Proof: Fix constants $\kappa > 0$ and $\varrho > 0$ so that $\sec_M \leq \kappa$ on $C_{t_{i-1}}$, the injectivity radius at each $p \in C_{t_{i-1}}$ is at least 3ϱ , and $3\varrho \leq D_\kappa = \text{diam}(\mathbb{M}_\kappa^2)$. Put $D := \text{diam}(C_{t_{i-1}})$. Given $\epsilon > 0$, choose $\delta > 0$ so that $\frac{D}{\epsilon}\delta < \varrho$. Now suppose that $t_{i-1} \leq r < s \leq t_i - \epsilon$ and $s - r \leq \delta$. Let $p, q \in C_r \setminus C_s$ with $d(p, q) < \varrho$, and choose nearest points $x, y \in C_s$, thus $d(p, x) = d(p, C_s)$ and $d(q, y) = d(q, C_s)$. Assume that $x \neq y$. By Lemma 7.8,

$$d(p, x) \leq \frac{\text{diam}(C_s)}{t_i - s}(s - r) \leq \frac{D}{\epsilon}(s - r) \leq \frac{D}{\epsilon}\delta < \varrho$$

and, likewise, $d(q, y) \leq \frac{D}{\epsilon}(s - r) < \varrho$. Thus $d(q, x) < 2\varrho$ and $d(x, y) < 3\varrho$. By the choice of ϱ there are unique segments pq, px, qx, qy, xy (possibly $d(p, q) = 0$), and the triples (p, q, x) and (q, x, y) have perimeter $< 6\varrho \leq 2D_\kappa$. Furthermore, $pq \cup px \cup qx \subset B(x, 2\varrho)$ and $qx \cup qy \cup xy \subset B(x, 3\varrho) \cap B(y, 3\varrho)$. As $xy \subset C_s$, a variation argument shows that $\angle_x(p, y) \geq \pi/2$ and $\angle_y(q, x) \geq \frac{\pi}{2}$. Consider points $\bar{p}, \bar{q}, \bar{x}, \bar{y} \in \mathbb{M}_\kappa^2$ such that $(\bar{p}, \bar{q}, \bar{x})$ and $(\bar{q}, \bar{x}, \bar{y})$ are comparison triples for (p, q, x) and (q, x, y) , respectively, and $\bar{p}, \bar{y}$ do not lie in the same connected component of $\mathbb{M}_\kappa^2 \setminus l_{\bar{q}, \bar{x}}$, where $l_{\bar{q}, \bar{x}}$ denotes the great circle through $\bar{q}, \bar{x}$. A similar application of Corollary 3.21 as in the first part of the proof of Theorem 5.12 shows that

$$\angle_{\bar{x}}(\bar{p}, \bar{q}) + \angle_{\bar{x}}(\bar{q}, \bar{y}) \geq \angle_x(p, q) + \angle_x(q, y) \geq \angle_x(p, y) \geq \frac{\pi}{2}$$

and $\angle_{\bar{y}}(\bar{q}, \bar{x}) \geq \angle_y(q, x) \geq \frac{\pi}{2}$. Now let $b \in (\frac{D}{\epsilon}(s - r), \frac{1}{2}D_\kappa)$, and suppose in addition that $d(p, q) = d(\bar{p}, \bar{q})$ is sufficiently small, so that the segment $\bar{p}\bar{q}$ lies in the closed b -neighborhood N_b of the great circle l through $\bar{x}, \bar{y}$. Let $\bar{\psi}: N_b \rightarrow l$ denote the nearest point retraction. The norm of the differential of $\bar{\psi}$ is everywhere less than or equal to $\lambda := \cos(\sqrt{\kappa}b)^{-1}$, hence $\bar{\psi}$ maps $\bar{p}\bar{q}$ to a subsegment of l of length at most $\lambda d(\bar{p}, \bar{q})$. Since $\angle_{\bar{x}}(\bar{p}, \bar{y}), \angle_{\bar{y}}(\bar{q}, \bar{x}) \geq \frac{\pi}{2}$, this subsegment contains $\bar{x}\bar{y}$, so

$$d(x, y) = d(\bar{x}, \bar{y}) \leq \lambda d(\bar{p}, \bar{q}) = \lambda d(p, q).$$

This shows at once that there is a unique nearest point retraction $\psi: C_r \rightarrow C_s$ and that $d(\psi(p), \psi(q)) \leq \lambda d(p, q)$ whenever $b \in (\frac{D}{\epsilon}(s - r), \frac{1}{2}D_\kappa)$ and $p, q \in C_r$ are sufficiently close to each other. Since any pair of points $p', q' \in C_r$ is connected by a segment $p'q' \subset C_r$, it follows that ψ is c^{-1} -Lipschitz for $c := \cos(\sqrt{\kappa}\frac{D}{\epsilon}(s - r))$. $\square$

Theorem 7.10 (Sharafutdinov) *There exists a 1-Lipschitz retraction $\phi: M \rightarrow \Sigma$.*

Proof: As said earlier, it suffices to show that for any fixed $i \leq k$ there is a 1-Lipschitz retraction $\phi_i: C_{t_{i-1}} \rightarrow C_{t_i}$. To this end, it is enough to construct a 1-Lipschitz retraction $\phi_{r,s}: C_r \rightarrow C_s$ for $r := t_{i-1}$ and every $s \in (t_{i-1}, t_i)$. Given such an s , set $\epsilon := t_i - s$, and let $\delta, a > 0$ be the respective constants provided by Proposition 7.9. Choose an integer $n \geq 1$ such that $\frac{1}{n}(s - r) \leq \delta$,

and subdivide $[r, s]$ into n subintervals $[r_0, r_1], \dots, [r_{n-1}, r_n]$ of length $\frac{1}{n}(s - r)$. The corresponding nearest point retractions $C_{r_{j-1}} \rightarrow C_{r_j}$ for $j = 1, \dots, n$ are c_n^{-1} -Lipschitz with $c_n := \cos(\frac{a}{n}(s - r))$. Taking their concatenation we obtain a c_n^{-n} -Lipschitz retraction $\psi_n: C_r \rightarrow C_s$. Now

$$c_n^{-1} = \cos\left(\frac{a}{n}(s - r)\right)^{-1} \leq 1 + \frac{c'}{n^2}$$

for some $c' > 0$ independent of n , and $c_n^{-n} \leq \left(1 + \frac{c'}{n^2}\right)^n \rightarrow 1$ as $n \rightarrow \infty$. It follows that some subsequence of (ψ_n) converges to a 1-Lipschitz retraction $\phi_{r,s}: C_r \rightarrow C_s$. $\square$

A refinement of this argument shows that there is a homotopy $H: M \times [0, 1] \rightarrow M$ from id_M to ϕ with $H(p, t) = p$ for all $(p, t) \in \Sigma \times [0, 1]$ (that is, H is a strong deformation retraction from M onto Σ). This can further be used to show that any two souls of M are isometric to each other; see [Yi1988] for details.

7.4 Proof of the Cheeger–Gromoll conjecture

Using the Sharafutdinov retraction and the Rauch–Berger comparison theorem (Theorem 3.26), Perelman [Pe1994] showed that the soul is a singleton already when $\sec(P) > 0$ for all planes $P \subset TM_q$ at some point $q \in M$, thus resolving the *Cheeger–Gromoll soul conjecture*.

Proposition 7.11 (Perelman) *Let $\phi: M \rightarrow \Sigma$ be any 1-Lipschitz retraction. For every unit vector $u \in T\Sigma^\perp$, let $\varrho_u: [0, \infty) \rightarrow M$ denote the geodesic with initial vector u . Then*

$$\phi(\varrho_u(t)) = \varrho_u(0)$$

for all u and $t \geq 0$. Furthermore, for every unit speed geodesic $\sigma: I \rightarrow \Sigma$ (I any interval) and every parallel unit vector field $N: I \rightarrow T\Sigma^\perp$ along σ , the map

$$F: I \times [0, \infty) \rightarrow M, \quad F(s, t) := \varrho_{N(s)}(t),$$

is a totally geodesic immersion that is isometric with respect to the standard inner product on $\mathbb{R}^2$, and if σ is minimizing, then so is each geodesic $\sigma_t := F(\cdot, t)$.

Proof: Define $f: [0, \infty) \rightarrow \mathbb{R}$ by

$$f(t) := \max\{d(\phi(\varrho_u(t)), p) : p \in \Sigma, u \in T\Sigma_p^\perp, |u| = 1\}.$$

Note that $f \geq 0 = f(0)$ and f is 1-Lipschitz: if $t, t' \geq 0$, and p and u are chosen such that $d(\phi(\varrho_u(t)), p) = f(t)$, then

$$\begin{aligned} f(t) - f(t') &\leq d(\phi(\varrho_u(t)), p) - d(\phi(\varrho_u(t')), p) \leq d(\phi(\varrho_u(t)), \phi(\varrho_u(t'))) \\ &\leq d(\varrho_u(t), \varrho_u(t')) \leq |t - t'|. \end{aligned}$$

We want to prove that f is identically zero. Consider the following two assertions for $a \geq 0$:

- (1)_a $f|_{[0,a]} \equiv 0$;
- (2)_a for every minimizing unit speed geodesic $\sigma: I \rightarrow \Sigma$ on some interval $I = [0, l]$ and every parallel unit vector field $N: I \rightarrow T\Sigma^\perp$ along σ , the map $F|_{I \times [0,a]}$ (where F is defined as above) is a totally geodesic isometric immersion, and $\sigma_t = F(\cdot, t)$ is minimizing for every $t \in [0, a]$.

The set A of all $a \geq 0$ such that (1)_a and (2)_a hold is closed and connected and contains 0. Hence, to prove the proposition, it suffices to show that for every $a \in A$ there is a $b > a$ contained in A .

Given $a \in A$ (possibly $a = 0$), we choose $b > a$ such that

- (i) $b - a$ is (strictly) smaller than the injectivity radius of Σ , and
- (ii) for every unit vector $u \in T\Sigma^\perp$, no normal Jacobi field Y along ϱ_u with $|Y(a)| = 1$ and $Y'(a) = 0$ has a zero in (a, b) .

Now fix $r \in (a, b]$ for the moment, and choose a pair (p, u) with $u \in T\Sigma_p^\perp$ such that $f(r) = d(\bar{p}, p)$, where $\bar{p} := \phi(\varrho_u(r))$. By (1)_a, $p = \phi(\varrho_u(a))$, thus

$$d(\bar{p}, p) \leq d(\varrho_u(r), \varrho_u(a)) \leq r - a \leq b - a.$$

Hence, by (i), there exists an $l > 0$ such that the segment $\bar{p}p$ can be prolongeded by a minimizing unit speed geodesic $\sigma: I \rightarrow \Sigma$ on $I = [0, l]$ from p to some point q such that $d(\bar{p}, p) + l = d(\bar{p}, q)$. Let N denote the parallel vector field along σ with $N(0) = u$, and set $v := N(l)$. Consider the corresponding map F . By (2)_a, σ_a is a geodesic of length l , and for every $s \in I$, the normal Jacobi field $t \mapsto \sigma'_t(s)$ along $\varrho_{N(s)}$ is parallel on $[0, a]$ or, in case $a = 0$, $\frac{D}{dt}|_{t=0} \sigma'_t(s) = \frac{D}{ds} N(s) = 0$. It now follows from (ii) and Theorem 3.26 that $|\sigma'_r(s)| \leq 1$ for all $s \in I$, thus $L(\sigma_r) \leq l$. Furthermore, an analysis of the proof shows that $L(\sigma_r) = l$ if and only if $F|_{I \times [a,r]}$ is a totally geodesic isometric immersion. Let $\bar{q} := \phi(\varrho_v(r))$. Then

$$d(\bar{p}, \bar{q}) \leq d(\varrho_u(r), \varrho_v(r)) \leq L(\sigma_r) \leq l; \quad (1)$$

on the other hand, by the choice of p and q and the definition of f ,

$$f(r) + l = d(\bar{p}, q) \leq d(\bar{p}, \bar{q}) + d(\bar{q}, q) \leq d(\bar{p}, \bar{q}) + f(r). \quad (2)$$

So all inequalities in (1) and (2) must in fact be equalities. In particular $F|_{I \times [a,r]}$ is a totally geodesic isometric immersion, and σ_r has minimal length $d(\varrho_u(r), \varrho_v(r))$. Now, for $r' \in (a, r)$ and $\bar{q}' := \phi(\varrho_v(r'))$,

$$d(\bar{p}, \bar{q}') \leq d(\varrho_u(r), \varrho_v(r')) \leq \sqrt{l^2 + (r - r')^2} \leq l + \frac{1}{2l}(r - r')^2$$

and $f(r) + l \leq d(\bar{p}, \bar{q}') + f(r')$, by (2) with $\bar{q}'$ in place of $\bar{q}$. Hence

$$\limsup_{r' \rightarrow r-} \frac{f(r) - f(r')}{r - r'} \leq 0.$$

This last inequality holds for every $r \in (a, b]$. As f is 1-Lipschitz and therefore absolutely continuous, and $f|_{[0,a]} \equiv 0 \leq f(r)$, an integration yields (1)_b. To prove (2)_b, let now $\sigma: I \rightarrow \Sigma$ be an arbitrary minimizing unit speed geodesic on some interval $I = [0, l]$, and put $p := \sigma(0)$ and $q := \sigma(l)$. Repeating a part of the above argument for $r \in (a, b]$, with $f(r) = 0$ and $l = d(p, q) = d(\bar{p}, \bar{q})$ in place of (2), we conclude that $F|_{I \times [a, r]}$ is a totally geodesic isometric immersion and σ_r is minimizing. Together with (2)_a, this gives (2)_b. $\square$

Theorem 7.12 (Perelman 1994) *If at some point $q \in M$ all sectional curvatures are positive, then every soul of M consists of a single point and hence M is diffeomorphic to $\mathbb{R}^m$.*

Proof: Suppose to the contrary that the soul Σ is not a singleton. Then it follows from the second part of Proposition 7.11 that for every $p \in \Sigma$ there is a plane $P \subset TM_p$ with $\sec(P) = 0$. Hence, the point q is not in Σ . Let $l > 0$ denote the distance from q to Σ , and choose a unit speed geodesic $\varrho: [0, l] \rightarrow M$ from Σ to q . Then $\varrho'(0)$ is normal, and since $\dim(\Sigma) \geq 1$, it follows again from the proposition that there exists a flat rectangle adjacent to ϱ . In particular, $\sec(P) = 0$ for some plane $P \subset TM_q$. $\square$

As another direct application of Proposition 7.11, it can be shown that any 1-Lipschitz retraction $\phi: M \rightarrow \Sigma$ is in fact a C^1 Riemannian submersion; see [Pe1994]. The regularity was later improved to C^2 and C^∞ in [Gui2000] and [Wi2007]. In particular, for every $p \in \Sigma$, the fiber $\phi^{-1}\{p\}$ is a smooth submanifold of M .

Chapter 8

Minimal entropy and Mostow's rigidity theorem

The main purpose of this chapter is to give a complete account of the entropy rigidity theorem of Besson, Courtois, and Gallot [BesCG1995] in the context of negative curvature, along the lines of the exposition in [BesCG1996]. The volume entropy of a closed Riemannian manifold is defined as the exponential growth rate of the volume of balls in the universal cover. The minimality of this quantity characterizes the locally symmetric spaces of dimension at least three among all homotopy equivalent manifolds with negative curvature and equal volume (see Theorem 8.22). This yields, as a direct consequence, Mostow's rigidity theorem [Mos1973] for rank one locally symmetric spaces (Theorem 8.23). The proof is, to some extent, constructive and uses an appealing barycenter method. We further discuss some closely related topics such as the construction of visual metrics and conformal structures, due to Bourdon [Bou1995], on the boundary at infinity of a space of negative curvature.

8.1 The boundary at infinity

Recall that a *ray* ϱ in a Riemannian manifold or a metric space X is a minimizing unit speed geodesic $\varrho: [0, \infty) \rightarrow X$.

Definition 8.1 Let X be a CAT(0) space. We call two rays ϱ, σ in X *asymptotic*, and we write $\varrho \sim \sigma$, if the convex function $s \mapsto d(\varrho(s), \sigma(s))$ is bounded. This defines an equivalence relation on the collection of all rays. The *boundary at infinity* (or *ideal boundary* or *geodesic boundary*) X_∞ of X is the set of equivalence classes. The class of ϱ is denoted ϱ_∞ . Intuitively, $\varrho_\infty = \xi \in X_\infty$ may be regarded as an infinitely remote endpoint of ϱ , and we may say that ϱ is *asymptotic to* ξ or *a ray from* $p = \varrho(0)$ *to* ξ . Similarly, if $c: \mathbb{R} \rightarrow X$ is a unit speed geodesic, then

$c_{-\infty}$ and c_{∞} will denote the classes of the backward and forward ray, respectively, determined by c .

Lemma 8.2 *Let X be a CAT(0) space, and let ϱ be a ray in X . Then for every $q \in X$ there exists a unique ray σ such that $\sigma(0) = q$ and σ is asymptotic to ϱ .*

Proof: Put $p := \varrho(0)$, and let $q \in X$. For every $t \geq 0$, let $\sigma_t: [0, l_t] \rightarrow X$ be the unit speed geodesic from q to $\varrho(t)$. Let $0 < t < t'$, and put $\lambda := \frac{t}{t'}$ and $x := \sigma_{t'}(\lambda l_{t'})$. By (Busemann) convexity,

$$d(\sigma_t(l_t), x) = d(\varrho(\lambda t'), x) \leq (1 - \lambda) d(p, q) \leq d(p, q).$$

For $l := \min\{l_t, \lambda l_{t'}\}$, it follows from the triangle inequality that $d(\sigma_t(l), \sigma_{t'}(l)) \leq 2 d(\sigma_t(l_t), x)$. Hence, for every $s \in [0, l]$, again by convexity,

$$d(\sigma_t(s), \sigma_{t'}(s)) \leq \frac{s}{l} d(\sigma_t(l), \sigma_{t'}(l)) \leq \frac{2s}{l} d(p, q).$$

Thus $(\sigma_t(s))_{l_t \geq s}$ is a Cauchy family for every $s \geq 0$, and the geodesics σ_t converge uniformly on compact sets to a ray σ issuing from q . To see that $\sigma \sim \varrho$, note that

$$d(\varrho(l_t), \sigma_t(l_t)) = |l_t - t| = |d(q, \varrho(t)) - d(p, \varrho(t))| \leq d(p, q),$$

thus $d(\varrho(s), \sigma_t(s)) \leq d(p, q)$ for all $s \in [0, l_t]$, and therefore $d(\varrho(s), \sigma(s)) \leq d(p, q)$ for all $s \geq 0$. It is clear that σ is unique. $\square$

Let now $o \in X$ be a fixed base point. To define a topology on

$$\bar{X} := X \cup X_{\infty},$$

we identify every point $x \in X$ with the map $\varrho_{ox}: [0, \infty) \rightarrow X$ such that $\varrho_{ox}|_{[0, d(o, x)]}$ is the unit speed geodesic from o to x and $\varrho(s) = x$ for all $s \geq d(o, x)$, and every boundary point $\xi \in X_{\infty}$ with the ray $\varrho_{o\xi}$ from o to ξ . We endow the set of all such eventually constant or usual rays issuing from o with the topology of uniform convergence on compact sets (in other words, the compact-open topology). This provides $\bar{X}$ with a (metrizable) topology. If X is an n -dimensional Hadamard manifold, then $\bar{X}$ is homeomorphic to a closed ball in $\mathbb{R}^n$. More generally, for a proper CAT(0) space X , it follows from the Arzelà–Ascoli theorem that $\bar{X}$ is still compact. The topology of $\bar{X}$ is also called the *cone topology*. A basis is given by the open balls in X and the “cones”

$$U_o(\xi, r, \epsilon) := \{\eta \in \bar{X} \setminus \bar{B}(o, r) : d(\varrho_{o\xi}(r), \varrho_{o\eta}(r)) < \epsilon\}$$

for $\xi \in X_{\infty}$ and $r, \epsilon > 0$. Similar arguments as in the proof of Lemma 8.2 show that the topology does not depend on the choice of o . Some general references are Section 2 in [EbON1973], Chapter II in [Ba1995], and Chapter II.8 in [BrH1999].

We remark further that every isometry $\gamma: X \rightarrow X$ induces homeomorphisms

$$\gamma_\infty: X_\infty \rightarrow X_\infty, \quad \bar{\gamma}: \bar{X} \rightarrow \bar{X},$$

where $\gamma_\infty(\varrho_\infty) = (\gamma \circ \varrho)_\infty$ for every ray $\varrho: [0, \infty) \rightarrow X$.

Next, we resume the discussion of Busemann functions from Section 7.1. Recall that the function $b_\varrho: X \rightarrow \mathbb{R}$ associated with a ray ϱ is defined by

$$b_\varrho(x) := \lim_{t \rightarrow \infty} d(x, \varrho(t)) - t.$$

Note that for fixed t , the distance function $d(\cdot, \varrho(t))$ is convex, so b_ϱ is convex, and therefore horoballs are convex. In a Hadamard manifold, Busemann functions are at least C^2 (see [HeIH1977]). The following result complements Lemma 7.1.

Proposition 8.3 *Let X be a CAT(0) space. If ϱ, σ are two asymptotic rays in X , then*

$$b_\varrho(x) - b_\sigma(x) = b_\varrho(\sigma(0)) = -b_\sigma(\varrho(0))$$

for all $x \in X$. In particular, horoballs of ϱ depend only on the class ϱ_∞ .

Proof: It follows as in the proof of Lemma 7.1 that $b_\sigma(x) \geq b_\varrho(x) - b_\varrho(\sigma(0))$. To prove the reverse inequality, let $t > 0$, put $a := d(x, \sigma(0))$, $b := d(\varrho(t), \sigma(t))$, and

$$r := d(x, \varrho(t)), \quad s := d(x, \sigma(t)), \quad u := d(\sigma(0), \varrho(t)).$$

By Proposition 5.23, $s^2 + u^2 \leq r^2 + t^2 + a^2 + b^2$ (where $t = d(\sigma(0), \sigma(t))$), thus

$$s - t \leq (r - u) \frac{r + u}{s + t} + \frac{a^2 + b^2}{s + t}.$$

Note that $|r - s| \leq b$, $|u - t| \leq b$, and $b \leq d(\varrho(0), \sigma(0))$ by convexity. Hence, as $t \rightarrow \infty$, the two fractions converge to 1 and 0, respectively, and writing $r - u$ as $(r - t) - (u - t)$ we conclude that $b_\sigma(x) \leq b_\varrho(x) - b_\varrho(\sigma(0))$. $\square$

Proposition 8.3 shows that $b_\varrho(x) - b_\varrho(y) = b_\sigma(x) - b_\sigma(y)$ whenever $\varrho \sim \sigma$ and $x, y \in X$. Thus, for every $\xi \in X_\infty$, there is a well-defined function $B_\xi: X \times X \rightarrow \mathbb{R}$, called the *Busemann cocycle* of ξ , such that

$$B_\xi(x, y) = b_\varrho(x) - b_\varrho(y) = \lim_{t \rightarrow \infty} (d(x, \varrho(t)) - d(y, \varrho(t)))$$

for every ray ϱ in the class ξ . Note that

$$B_\xi(y, x) = -B_\xi(x, y), \quad B_\xi(x, y) + B_\xi(y, z) = B_\xi(x, z),$$

and $|B_\xi(x, y)| \leq d(x, y)$ for all $x, y, z \in X$.

Proposition 8.4 *The function $B: \bar{X} \times X \times X \rightarrow \mathbb{R}$,*

$$B(\xi, x, y) := \begin{cases} B_\xi(x, y) & \text{if } \xi \in X_\infty, \\ d(x, \xi) - d(y, \xi) & \text{if } \xi \in X, \end{cases}$$

is continuous.

Proof: Fix $o \in X$. Since $B(\xi, x, y) = B(\xi, x, o) - B(\xi, y, o)$ and $B(\xi, \cdot, o)$ is 1-Lipschitz, it suffices to verify that $B(\cdot, x, o)$ is continuous on $\bar{X}$ for any $x \in X$. For this, let ϱ be the ray from o to a boundary point $\xi \in X_\infty$, and let σ be the eventually constant or usual ray from o to a point $\eta \in \bar{X}$. Comparison with $\mathbb{R}^2$ shows that for every $\epsilon > 0$ there is a constant $r > 0$, depending otherwise only on $d(o, x)$, such that if z' is a point on a segment oz and $d(o, z') \geq r$, then

$$d(x, z) - d(o, z) \leq d(x, z') - d(o, z') \leq d(x, z) - d(o, z) + \epsilon.$$

Hence, if $\eta \in \bar{X}$ is in a sufficiently small neighborhood of ξ , such that $\eta \notin \bar{B}(o, r)$ and $d(\varrho(r), \sigma(r)) < \epsilon$, then

$$B(\eta, x, o) \leq d(x, \sigma(r)) - r < d(x, \varrho(r)) - r + \epsilon \leq B(\xi, x, o) + 2\epsilon$$

(where $r = d(o, \sigma(r)) = d(o, \varrho(r))$), and likewise $B(\xi, x, o) < B(\eta, x, o) + 2\epsilon$. $\square$

The following result complements the discussion of isometries in Section 4.5 (see also Corollary 5.26).

Theorem 8.5 *Suppose that X is a proper CAT(0) space, and γ is a parabolic isometry of X . Then $\bar{\gamma}: \bar{X} \rightarrow \bar{X}$ fixes at least one point $\xi \in X_\infty$, and γ leaves invariant all horoballs centered at ξ .*

Proof: The displacement function $d_\gamma: X \rightarrow [0, \infty)$, $d_\gamma(x) := d(x, \gamma(x))$, is convex, γ -invariant, and does not attain its infimum $|\gamma|$ (compare 4.21–4.23). Hence, for every $\lambda > |\gamma|$, the closed sublevel set $C_\lambda := \{x \in X : d_\gamma(x) \leq \lambda\} \neq \emptyset$ is convex and γ -invariant. Fix a basepoint $p \in X$, and let p_λ be the point in C_λ nearest to p . Note that for every $R > 0$ there is an $\epsilon > 0$ such that $d_\gamma \geq |\gamma| + \epsilon$ on $\bar{B}(p, R)$, thus $d(p, p_\lambda) > R$ for $\lambda < |\gamma| + \epsilon$. This shows that $d(p, p_\lambda) \rightarrow \infty$ as $\lambda \rightarrow |\gamma|$. Since $\bar{X}$ is compact, there exists a sequence $\lambda_k \rightarrow |\gamma|$ such that p_{λ_k} converges to some point $\xi \in X_\infty$ as $k \rightarrow \infty$. Let ϱ be the ray from p to ξ . If k is sufficiently large, then $d_\gamma(p_{\lambda_k}) \leq \lambda_k \leq d_\gamma(p)$, and since $p_{\lambda_k} \rightarrow \xi$ and d_γ is convex, it follows that $d_\gamma(\varrho(t)) \leq d_\gamma(p)$ for all $t \geq 0$. Hence, ϱ is asymptotic to $\gamma \circ \varrho$, and so $\bar{\gamma}$ fixes ξ .

To show that the Busemann function b_ϱ is γ -invariant, that is, $b_\varrho \circ \gamma = b_\varrho$, consider the functions $f_\lambda, h_\lambda: X \rightarrow \mathbb{R}$,

$$f_\lambda(q) := d(q, p_\lambda) - d(p, p_\lambda), \quad h_\lambda(q) := d(q, q_\lambda) - d(p, p_\lambda),$$

where q_λ denotes the point in C_λ closest to q . By Proposition 8.4, as $k \rightarrow \infty$,

$$f_{\lambda_k}(q) = B(p_{\lambda_k}, q, p) \rightarrow B(\xi, q, p) = b_\varrho(q). \quad (1)$$

Since C_λ is γ -invariant, so is h_λ . Put $a := d(p, q)$, $b := d(p_\lambda, q_\lambda)$, and

$$r := d(p, p_\lambda), \quad s := d(p, q_\lambda), \quad t := d(q, q_\lambda), \quad u := d(q, p_\lambda).$$

By Proposition 5.23, $s^2 + u^2 \leq r^2 + t^2 + a^2 + b^2$, furthermore $r \leq s$ and $b \leq a$ (Proposition 5.24). Hence, $u^2 - r^2 \leq t^2 - r^2 + 2a^2$ and

$$u - r \leq (t - r) \frac{t + r}{u + r} + \frac{2a^2}{u + r}.$$

As $\lambda \rightarrow |\gamma|$, the two fractions converge to 1 and 0, respectively. Since $u - r = f_\lambda(q)$ and $t - r = h_\lambda(q) \leq f_\lambda(q)$, we conclude from (1) that $h_{\lambda_k}(q) \rightarrow b_\varrho(q)$ for all $q \in X$. Thus, since h_λ is γ -invariant, so is b_ϱ . $\square$

8.2 Visual metrics

We now assume that X is a Hadamard manifold with sectional curvature at most -1 or, more generally, a $\text{CAT}(-1)$ space. Our next goal is to equip the boundary at infinity of X with a family $(d_o)_{o \in X}$ of so-called *visual metrics* compatible with the subspace topology of $X_\infty \subset \bar{X}$.

Given a geodesic triangle with vertices o, x, y in X , there is a unique triple of points $u \in ox$, $v \in oy$, and $w \in xy$ such that any two of them are at equal distance to the common vertex of the respective sides, thus $d(u, o) = d(v, o)$, $d(v, y) = d(w, y)$, and $d(w, x) = d(u, x)$. Following [BuyS2007] we call these the *equiradial points* of the triangle. The distance of u and v to o is explicitly given by the quantity

$$(x | y)_o := \frac{1}{2} (d(o, x) + d(o, y) - d(x, y)),$$

called the *Gromov product* of x and y with respect to o due to its central role in [Gro1987]. If $\tilde{o} \in X$ is another base point, then

$$|(x | y)_o - (x | y)_{\tilde{o}}| = \frac{1}{2} |d(o, x) + d(o, y) - d(\tilde{o}, x) - d(\tilde{o}, y)| \leq d(o, \tilde{o}).$$

Furthermore, the Gromov product possesses the following monotonicity property: if $x' \in ox$ and $y' \in oy$, then

$$(x' | y')_o \leq (x | y)_o,$$

because $2(x | y)_o - 2(x' | y')_o = d(x, x') + d(y, y') - d(x, y) + d(x', y') \geq 0$ by the triangle inequality. Thus, for two points $\xi, \eta \in X_\infty$ represented by the rays ϱ, σ issuing from o , we can define

$$(\xi | \eta)_o := \lim_{s, t \rightarrow \infty} (\varrho(s) | \sigma(t))_o \in [0, \infty].$$

As the notation suggests, the limit does not depend on the choice of representatives. To verify this, suppose that $\tilde{\varrho} \sim \varrho$ and $\tilde{\sigma} \sim \sigma$. Comparison with H^2 shows that if two asymptotic rays in X are properly synchronized, so that the Busemann functions agree, then the distance between them is not only bounded but tends to zero. That is, there exist $\alpha, \beta \in \mathbb{R}$ such that $d(\tilde{\varrho}(s), \varrho(s+\alpha)) \rightarrow 0$ and $d(\tilde{\sigma}(t), \sigma(t+\beta)) \rightarrow 0$ as $s, t \rightarrow \infty$. In fact, by Proposition 8.3, $\alpha = b_{\tilde{\varrho}}(o)$ and $\beta = b_{\tilde{\sigma}}(o)$. It now follows that $\lim_{s,t \rightarrow \infty} (\tilde{\varrho}(s) | \tilde{\sigma}(t))_o = (\xi | \eta)_o$.

Lemma 8.6 *There exists a constant $\delta > 0$ such that if X is a CAT(-1) space and $\xi, \eta \in X_\infty$ are two distinct points represented by the rays ϱ, σ issuing from $o \in X$, then*

$$(\xi | \eta)_o \leq \sup\{r \geq 0 : d(\varrho(r), \sigma(r)) \leq \delta\} < \infty.$$

If X is proper in addition, then there exists a unit speed geodesic $c : \mathbb{R} \rightarrow X$ such that $c_{-\infty} = \xi$, $c_\infty = \eta$, and

$$(\xi | \eta)_o \leq d(o, \text{im}(c)) \leq (\xi | \eta)_o + \delta.$$

Note that c is unique up to reparametrization (compare Proposition 4.20).

Proof: Consider a geodesic triangle in X with vertices o, x, y . By comparison with H^2 , there exists a constant $\delta > 0$ such that the diameter of the set of equiradial points is at most δ , regardless of the size of the triangle. Hence, if $u \in ox$ is the point at distance $(x | y)_o$ from o , and $w \in xy$ is the point with $d(w, x) = d(u, x)$, then

$$d(o, xy) \leq d(o, w) \leq d(o, u) + d(u, w) \leq (x | y)_o + \delta.$$

Moreover, if $z \in xw$, then $d(z, x) \leq d(w, x) = d(u, x)$ and thus

$$d(o, z) \geq d(o, x) - d(z, x) \geq d(o, x) - d(u, x) = (x | y)_o.$$

So by symmetry, $d(o, xy) \geq (x | y)_o$.

Let now $\xi, \eta \in X_\infty$ be two distinct points represented by the rays ϱ, σ issuing from o . Since $\varrho \neq \sigma$, the supremum $R := \sup\{r : d(\varrho(r), \sigma(r)) \leq \delta\}$ is finite. Let $s > 0$, and consider the triangle with vertices $o, \varrho(s), \sigma(s)$. For $r := (\varrho(s) | \sigma(s))_o$, we have $d(\varrho(r), \sigma(r)) \leq \delta$ by the choice of δ , and so

$$(\varrho(s) | \sigma(s))_o = r \leq R.$$

This yields the first assertion of the lemma. For the second, if X is proper, put $l_s := d(\varrho(s), \sigma(s))$, and let $c_s : [-\frac{1}{2}l_s, \frac{1}{2}l_s] \rightarrow X$ be the unit speed geodesic from $\varrho(s)$ to $\sigma(s)$. By the above inequalities, $d(o, c_s(0)) \leq R + \delta$, and the desired geodesic c is obtained as the pointwise limit of some subsequence (c_{s_i}) . $\square$

The following result is due to Bourdon [Bou1995]. We give a somewhat different proof suggested by V. Schroeder.

Proposition 8.7 (visual metrics) *Let X be a $\text{CAT}(-1)$ space. For $o \in X$, the function $d_o: X_\infty \times X_\infty \rightarrow [0, 2]$ defined by*

$$d_o(\xi, \eta) := 2e^{-(\xi | \eta)_o},$$

with the convention $e^{-\infty} := 0$, is a metric on X_∞ . If $\tilde{o}$ is another base point, then

$$\lambda^{-1} d_o(\xi, \eta) \leq d_{\tilde{o}}(\xi, \eta) = d_o(\xi, \eta) e^{(B_\xi(o, \tilde{o}) + B_\eta(o, \tilde{o}))/2} \leq \lambda d_o(\xi, \eta)$$

for all $\xi, \eta \in X_\infty$, where $\lambda := e^{d(o, \tilde{o})}$.

Proof: We shall use comparison with the Poincaré unit disk (U, d) . The hyperbolic distance $d(\bar{x}, \bar{y})$ between $\bar{x}, \bar{y} \in U$ is given by the formula

$$\sinh^2\left(\frac{1}{2}d(\bar{x}, \bar{y})\right) = \frac{\cosh(d(\bar{x}, \bar{y})) - 1}{2} = \frac{\|\bar{x} - \bar{y}\|^2}{(1 - \|\bar{x}\|^2)(1 - \|\bar{y}\|^2)},$$

where $\|\cdot\|$ is the Euclidean norm (see [BrH1999], p. 86). It follows readily that $\cosh^2\left(\frac{1}{2}d(\bar{x}, 0)\right) = (1 - \|\bar{x}\|^2)^{-1}$ and

$$\|\bar{x} - \bar{y}\| = \frac{\sinh\left(\frac{1}{2}d(\bar{x}, \bar{y})\right)}{\cosh\left(\frac{1}{2}d(\bar{x}, 0)\right) \cosh\left(\frac{1}{2}d(\bar{y}, 0)\right)}.$$

In particular, if $d(\bar{x}, 0) = d(\bar{y}, 0) = r$ and $d(\bar{x}, \bar{y}) = s$, then

$$\|\bar{x} - \bar{y}\| = f_r(s) := \frac{\sinh\left(\frac{s}{2}\right)}{\cosh^2\left(\frac{r}{2}\right)} = 2 \frac{e^{s/2} - e^{-s/2}}{e^r + 2 + e^{-r}}.$$

We now verify the triangle inequality for d_o . Let ϱ, σ, τ be three distinct rays in X emanating from o . Fix $r > 0$ for the moment, and put $x := \varrho(r)$, $y := \sigma(r)$, and $z := \tau(r)$. Choose $\bar{x}, \bar{y}, \bar{z}$ in (U, d) such that $\bar{x}, \bar{y}, 0$ and $\bar{y}, \bar{z}, 0$ are comparison triples for x, y, o and y, z, o , respectively, and the two hyperbolic triangles have disjoint interiors. Suppose first that $\angle_0(\bar{x}, \bar{y}) + \angle_0(\bar{y}, \bar{z}) < \pi$. Then the hyperbolic segment $\bar{x}\bar{z}$ intersects $0\bar{y}$ in a point $\bar{w}$. Let $w := \sigma(d(\bar{w}, 0))$ be the corresponding point in X . By comparison,

$$d(\bar{x}, \bar{z}) = d(\bar{x}, \bar{w}) + d(\bar{w}, \bar{z}) \geq d(x, w) + d(w, z) \geq d(x, z).$$

Hence, by the Euclidean triangle inequality for $\bar{x}, \bar{y}, \bar{z}$ and the monotonicity of the function f_r defined above,

$$f_r(d(\bar{x}, \bar{y})) + f_r(d(\bar{y}, \bar{z})) \geq f_r(d(\bar{x}, \bar{z})) \geq f_r(d(x, z)).$$

In the remaining case, when $\angle_0(\bar{x}, \bar{y}) + \angle_0(\bar{y}, \bar{z}) \geq \pi$, we choose a comparison triple $\tilde{x}, \tilde{z}, 0 \in U$ for x, z, o ; then

$$f_r(d(\bar{x}, \bar{y})) + f_r(d(\bar{y}, \bar{z})) \geq \|\bar{x}\| + \|\bar{z}\| = \|\tilde{x}\| + \|\tilde{z}\| \geq f_r(d(\tilde{x}, \tilde{z})).$$

We conclude that, in either case,

$$f_r(d(x, y)) + f_r(d(y, z)) \geq f_r(d(x, z)).$$

Recall that $x = \varrho(r)$, $y = \sigma(r)$, and $z = \tau(r)$. Letting $r \rightarrow \infty$, we get that

$$\begin{aligned} \lim_{r \rightarrow \infty} f_r(d(\varrho(r), \sigma(r))) &= \lim_{r \rightarrow \infty} 2e^{d(\varrho(r), \sigma(r))/2-r} \\ &= 2e^{-(\varrho_\infty | \sigma_\infty)_o} = d_o(\varrho_\infty, \sigma_\infty), \end{aligned}$$

and the inequality $d_o(\varrho_\infty, \sigma_\infty) + d_o(\sigma_\infty, \tau_\infty) \geq d_o(\varrho_\infty, \tau_\infty)$ follows.

Let now $\tilde{o} \in X$ be another base point. For two rays ϱ, σ as above, let $\tilde{\varrho}, \tilde{\sigma}$ be the rays emanating from $\tilde{o}$ with $\tilde{\varrho}_\infty = \xi := \varrho_\infty$ and $\tilde{\sigma}_\infty = \eta := \sigma_\infty$. Put $\alpha := b_{\tilde{\varrho}}(o) = B_\xi(o, \tilde{o})$ and $\beta := b_{\tilde{\sigma}}(o) = B_\eta(o, \tilde{o})$. Then $\lim_{r \rightarrow \infty} d(\tilde{\varrho}(r), \varrho(r + \alpha)) = 0$ and $\lim_{r \rightarrow \infty} d(\tilde{\sigma}(r), \sigma(r + \beta)) = 0$, and it follows that $(\xi | \eta)_{\tilde{o}} = (\xi | \eta)_o - \frac{1}{2}(\alpha + \beta)$. This yields the result. $\square$

The boundary map $\gamma_\infty: X_\infty \rightarrow X_\infty$ induced by an isometry $\gamma \in \text{Isom}(X)$ satisfies $(\gamma_\infty(\xi) | \gamma_\infty(\eta))_o = (\xi | \eta)_{\gamma^{-1}(o)}$ for all $\xi, \eta \in X_\infty$, hence

$$d_o(\gamma_\infty(\xi), \gamma_\infty(\eta)) = d_{\gamma^{-1}(o)}(\xi, \eta).$$

Proposition 8.7 then shows that γ_∞ is a λ -bi-Lipschitz homeomorphism with respect to d_o for $\lambda = e^{d(o, \gamma^{-1}(o))} = e^{d(\gamma(o), o)}$. A stronger property of γ_∞ is discussed next.

A map $f: (A, d) \rightarrow (B, d)$ between two metric spaces is called *conformal* if for every non-isolated point $a \in A$, the limit

$$\lim_{a' \rightarrow a} \frac{d(f(a), f(a'))}{d(a, a')},$$

called the *conformal factor* of f at a , exists and is (finite and) non-zero. Two metrics $d, \tilde{d}$ on A are conformal if the identity map $\text{id}: (A, d) \rightarrow (A, \tilde{d})$ is. A collection of pairwise conformal metrics is called a *conformal structure*. The following result is also due to Bourdon [Bou1995].

Proposition 8.8 (conformal structure) *Let X be a $\text{CAT}(-1)$ space. The family $(d_o)_{o \in X}$ is a conformal structure on X_∞ with conformal factors*

$$\lim_{\xi' \rightarrow \xi} \frac{d_{\tilde{o}}(\xi, \xi')}{d_o(\xi, \xi')} = e^{B_\xi(o, \tilde{o})}.$$

For an isometry $\gamma \in \text{Isom}(X)$, the induced map γ_∞ of (X_∞, d_o) is conformal with conformal factor $e^{B_\xi(o, \gamma^{-1}(o))}$ at ξ .

Proof: Let $o, \tilde{o} \in X$. We know from Proposition 8.4 that the function $\xi \mapsto B_\xi(o, \tilde{o})$ is continuous on X_∞ ; alternatively, this follows from the equation in Proposition 8.7 by considering a neighborhood $V \subset X_\infty$ of ξ and a fixed η outside of V . Now, for every non-isolated point $\xi \in X_\infty$,

$$\lim_{\xi' \rightarrow \xi} \frac{d_{\tilde{o}}(\xi, \xi')}{d_o(\xi, \xi')} = \lim_{\xi' \rightarrow \xi} e^{(B_\xi(o, \tilde{o}) + B_{\xi'}(o, \tilde{o}))/2} = e^{B_\xi(o, \tilde{o})}.$$

If $\gamma \in \text{Isom}(X)$, then $d_o(\gamma_\infty(\xi), \gamma_\infty(\xi')) = d_{\gamma^{-1}(o)}(\xi, \xi')$, as mentioned above, and the result follows. $\square$

8.3 Quasi-isometries

Let, for the moment, (X, d) and (Y, d) be arbitrary metric spaces. A (not necessarily continuous) map $f: X \rightarrow Y$ is called a (λ, α) -quasi-isometric embedding, for constants $\lambda \geq 1$ and $\alpha \geq 0$, if

$$\lambda^{-1}d(x, x') - \alpha \leq d(f(x), f(x')) \leq \lambda d(x, x') + \alpha$$

for all $x, x' \in X$, and f is called a *quasi-isometric embedding* if this holds for some such constants. If in addition $\sup\{d(y, f(X)) : y \in Y\} < \infty$, then f is called a *quasi-isometry*; then there also exists a quasi-isometry $g: Y \rightarrow X$ such that $g \circ f$ and $f \circ g$ are at finite uniform distance from the identity map on X and Y , respectively (a *rough inverse* of f). Two metric spaces are said to be *quasi-isometric* if there exists a quasi-isometry between them. The composition of two quasi-isometric embeddings or two quasi-isometries is again a quasi-isometric embedding or a quasi-isometry, respectively.

A main source of quasi-isometries is the following result, which complements Lemma 6.10.

Lemma 8.9 (orbit map) *Let X be a complete Riemannian manifold, and let $\Gamma \subset \text{Isom}(X)$ be a group of isometries acting properly discontinuously and cocompactly on X . For every word metric d_A on the (finitely generated) group Γ and every $x \in X$, the orbit map $(\Gamma, d_A) \rightarrow X, \gamma \mapsto \gamma(x)$, is a quasi-isometry.*

This holds more generally for any proper geodesic metric space X .

Proof: Given $x \in X$, choose $r > 0$ such that $\bigcup_{\gamma \in \Gamma} \bar{B}(\gamma(x), r) = X$. Lemma 6.10 (for $K = \bar{B}(\gamma(x), r)$ and $\delta = r$) shows that the set $A := \{a \in \Gamma : d(x, a(x)) < 3r\}$ is finite and generates Γ , and if $d(x, \gamma(x)) < kr$ for some integer $k \geq 1$, then $|\gamma|_A \leq k$. Now let $\gamma, \gamma' \in \Gamma$. If $k \geq 1$ is such that $(k-1)r \leq d(\gamma(x), \gamma'(x)) = d(x, \gamma^{-1}\gamma'(x)) < kr$, then

$$d_A(\gamma, \gamma') = |\gamma^{-1}\gamma'|_A \leq k \leq \frac{1}{r} d(\gamma(x), \gamma'(x)) + 1.$$

Conversely, suppose that $d_A(\gamma, \gamma') = k \geq 1$. There exist $a_1, \dots, a_k \in A$ such that $\gamma^{-1}\gamma' = a_1 \cdots a_k$. Put $\gamma_0 := \gamma$ and $\gamma_i := \gamma a_1 \cdots a_i$ for $i = 1, \dots, k$; then

$$d(\gamma(x), \gamma'(x)) \leq \sum_{i=1}^k d(\gamma_{i-1}(x), \gamma_i(x)) = \sum_{i=1}^k d(x, a_i(x)) < 3rk = 3r d_A(\gamma, \gamma').$$

Thus, for A as above, the orbit map $x \mapsto \gamma(x)$ is a quasi-isometry. Since any two word metrics are bi-Lipschitz equivalent to each other, the result follows. $\square$

The following characteristic property of spaces of negative curvature goes back to H. M. Morse [Mo1924] for the hyperbolic plane.

Proposition 8.10 (Morse lemma) *For all $\lambda \geq 1$ and $\alpha \geq 0$ there exists a constant $D \geq 0$ such that if X is a $\text{CAT}(-1)$ space and $\sigma: [0, l] \rightarrow X$ is a (λ, α) -quasi-isometric embedding, then the Hausdorff distance between $\text{im}(\sigma)$ and the segment pq connecting $p := \sigma(0)$ and $q := \sigma(l)$ is at most D .*

Proof: Let $r > 0$, to be specified later. Note that $d(\sigma(s), \sigma(s')) \leq \lambda + \alpha =: b$ whenever $|s - s'| \leq 1$. Hence, every image point of σ at distance $> r + b$ from pq is contained in some sequence $x_i := \sigma(t + i)$, for $i = 0, 1, \dots, k$, such that $d(x_i, pq) > r + b$ for $i = 1, \dots, k - 1$ and $d(x_0, pq), d(x_k, pq) \leq r + b$. Then

$$L := \sum_{i=1}^k d(x_{i-1}, x_i) \leq kb \leq b\lambda(d(x_0, x_k) + \alpha),$$

as $d(x_0, x_k) = d(\sigma(t), \sigma(t + k)) \geq \lambda^{-1}k - \alpha$. Let $\pi: X \rightarrow pq$ be the nearest point retraction, and notice that

$$d(x_0, x_k) \leq d(\pi(x_0), \pi(x_k)) + 2(r + b).$$

The segments $x_0x_1, \dots, x_{k-1}x_k$ form a curve of length L at distance $> r$ from pq . A comparison argument with H^2 then shows that

$$d(\pi(x_0), \pi(x_k)) \leq \cosh(r)^{-1}L.$$

(compare the proof of Proposition 7.9). We now choose r so that $b\lambda \cosh(r)^{-1} = \frac{1}{2}$. Altogether, this yields the bound $\frac{1}{2}L \leq b\lambda(2(r + b) + \alpha) =: C$, and we conclude that $\text{im}(\sigma)$ is within distance at most $r + b + C$ from pq .

Conversely, let $x \in pq$. Since $d(\pi(\sigma(s)), \pi(\sigma(s'))) \leq \lambda|s - s'| + \alpha$ for all $s, s' \in [0, l]$, there must exist a point $\pi(\sigma(t))$ at distance at most α from x . Then

$$d(x, \sigma(t)) \leq \alpha + d(\pi(\sigma(t)), \sigma(t)) \leq \alpha + r + b + C.$$

This completes the proof. $\square$

Among the many applications of the Morse lemma is the following key result.

Theorem 8.11 (induced boundary map) *Let X, Y be two Hadamard manifolds of curvature ≤ -1 , or two proper $\text{CAT}(-1)$ spaces, and suppose that $f: X \rightarrow Y$ is a (λ, α) -quasi-isometric embedding. Then f induces an embedding $f_\infty: X_\infty \rightarrow Y_\infty$ such that both f_∞ and its inverse on $\text{im}(f_\infty)$ are Hölder continuous with exponent λ^{-1} with respect to d_o and $d_{f(o)}$ for any base point $o \in X$. If f is a quasi-isometry, then f_∞ is a homeomorphism.*

Proof: To define the map f_∞ , let $\xi \in X_\infty$, and let ϱ be the ray from the base point o to ξ . By Proposition 8.10 there exists a constant $D = D(\lambda, \alpha)$ such that for every $t > 0$, the image of $f \circ \varrho|_{[0,t]}$ is at Hausdorff distance at most D from the segment connecting $f(o)$ and $f \circ \varrho(t)$. Let $\bar{\varrho}_t: [0, l_t] \rightarrow Y$ be the unit speed parametrization of this segment. Notice that

$$l_t = d(f(o), f \circ \varrho(t)) \geq \lambda^{-1}t - \alpha,$$

and the endpoint $\bar{\varrho}_t(l_t)$ is at distance $\leq D$ from $\text{im}(\bar{\varrho}_{t'})$ for any $t' > t$. By convexity, $(\bar{\varrho}_t(s))_{t \geq s}$ is a Cauchy family for every $s > 0$, and the geodesics $\bar{\varrho}_t$ converge uniformly on compact sets to a ray $\bar{\varrho}$ in Y as $t \rightarrow \infty$. The trace of $\bar{\varrho}$ is at Hausdorff distance $\leq D$ from $\text{im}(f \circ \varrho)$, and no other ray emanating from $f(o)$ remains at finite distance from this set. Now put $f_\infty(\xi) := \bar{\varrho}_\infty$. If σ is the ray from a different base point in X to ξ , then $s \mapsto d(f \circ \varrho(s), f \circ \sigma(s))$ is uniformly bounded, and the corresponding ray $\bar{\sigma}$ in Y is asymptotic to $\bar{\varrho}$; thus f_∞ is independent of the choice of o .

Now let $\xi, \eta \in X_\infty$ be two distinct points. By Lemma 8.6 there exists a unit speed geodesic $c: \mathbb{R} \rightarrow X$ from ξ to η . A similar argument as above, still relying on Proposition 8.10, but using that Y is proper, provides a unit speed geodesic $\bar{c}: \mathbb{R} \rightarrow Y$ such that $\text{im}(\bar{c})$ is at Hausdorff distance $\leq D$ from $\text{im}(f \circ c)$ and $\bar{c}$ is backward asymptotic to $\bar{\xi} := f_\infty(\xi)$ and forward asymptotic to $\bar{\eta} := f_\infty(\eta)$. Thus

$$\begin{aligned} \lambda^{-1}d(o, \text{im}(c)) - \alpha &\leq d(f(o), \text{im}(f \circ c)) \leq d(f(o), \text{im}(\bar{c})) + D, \\ d(f(o), \text{im}(\bar{c})) - D &\leq d(f(o), \text{im}(f \circ c)) \leq \lambda d(o, \text{im}(c)) + \alpha. \end{aligned}$$

Together with the inequalities in Lemma 8.6, for o and c as well as $f(o)$ and $\bar{c}$, this gives $\lambda^{-1}(\xi | \eta)_o - \alpha \leq (\bar{\xi} | \bar{\eta})_{f(o)} + \delta + D$, hence

$$d_{f(o)}(\bar{\xi}, \bar{\eta}) \leq e^{\delta+D+\alpha} d_o(\xi, \eta)^{1/\lambda},$$

and $(\bar{\xi} | \bar{\eta})_{f(o)} - D \leq \lambda(\xi | \eta)_o + \lambda\delta + \alpha$, hence

$$d_o(\xi, \eta) \leq e^{\delta+(D+\alpha)/\lambda} d_{f(o)}(\bar{\xi}, \bar{\eta})^{1/\lambda}.$$

If f is a quasi-isometry and $g: Y \rightarrow X$ is a rough inverse, then $f \circ g$ is at finite uniform distance from id_Y , and $f_\infty \circ g_\infty$ is the identity map on Y_∞ , thus f_∞ is surjective. $\square$

Another fact about quasi-isometries used later on is that every *continuous* such map between two Hadamard manifolds (of the same dimension) is surjective. This follows readily, via the exponential map, from the following lemma for $\mathbb{R}^n$ (compare Lemma 10.1 in [Mos1973] and Lemma 2.1 in [LaS1997]).

Lemma 8.12 *Let $R, r > 0$, and suppose that f is a continuous map from $\bar{B}_R := \{x \in \mathbb{R}^n : \|x\| \leq R\}$ to $\mathbb{R}^n$ such that $f(0) = 0$ and $\|f(x) - f(y)\| \geq r$ for every pair of points $x, y \in \bar{B}_R$ with $\|x\| + \|y\| = \|x - y\| = R$. Then $\bar{B}_r \subset \text{im}(f)$.*

Proof: Suppose to the contrary that $\bar{B}_r \not\subset \text{im}(f)$. Then f also omits a point in the open ball B_r , and by radially pushing out f we obtain a continuous map $\tilde{f}: \bar{B}_R \rightarrow \mathbb{R}^n \setminus B_r$ that agrees with f on $\partial\bar{B}_R$, because $\|f(x)\| \geq r$ for all $x \in \partial\bar{B}_R$. Define the map $g: \bar{B}_R \rightarrow S^{n-1}$ to the unit sphere by $g(x) := \tilde{f}(x)/\|\tilde{f}(x)\|$, and consider the homotopy $h: \partial\bar{B}_R \times [0, 1] \rightarrow S^{n-1}$,

$$h_t(x) := \frac{f(x - \frac{t}{2}x) - f(-\frac{t}{2}x)}{\|f(x - \frac{t}{2}x) - f(-\frac{t}{2}x)\|};$$

the denominator is non-zero by the assumption in the lemma. Notice that h_0 agrees with the restriction of g to $\partial\bar{B}_R$, and h_1 satisfies $h_1(-x) = -h_1(x)$ for all $x \in \partial\bar{B}_R$. By combining g and h and identifying the domain $\bar{B}_R \cup (\partial\bar{B}_R \times [0, 1])$ with a closed hemisphere of S^n , we get a continuous map from this hemisphere to S^{n-1} that maps pairs of antipodal points in the boundary to pairs of antipodal points. Central reflection then gives rise to a map from S^n to S^{n-1} preserving antipodes, in contradiction to the Borsuk–Ulam theorem. $\square$

8.4 Entropy and conformal densities

We now turn to the volume entropy and to the construction of a family $(\mu_x)_{x \in X}$ of measures on X_∞ due to Patterson [Pat1976] and Sullivan [Su1979].

For the following two preliminary results, $\pi: X \rightarrow M$ denotes the Riemannian universal cover of an arbitrary closed (that is, compact and connected) Riemannian manifold M with an infinite group $\Gamma \subset \text{Isom}(X)$ of deck transformations. For $o, x \in X$ and $r > 0$, let

$$N_o(x, r) := \#(\Gamma(o) \cap B(x, r))$$

denote the number of orbit points in the open ball $B(x, r)$.

Lemma 8.13 (entropy) *With this notation, the limits*

$$h(M) := \lim_{r \rightarrow \infty} \frac{1}{r} \log(\text{vol}(B(x, r))) = \lim_{r \rightarrow \infty} \frac{1}{r} \log(N_o(x, r)) \in [0, \infty)$$

exist and do not depend on the choice of x and o .

This defines the *volume entropy*, or briefly the *entropy*, of M . The result originates from [Ma1979] and is inspired by [Mi1968] and [Sv1955].

Proof: Choose $a > 0$ such that every $y \in X$ is at distance at most a from the orbit $\Gamma(o)$. Let $r, t > a$. Every point in $\Gamma(o) \cap B(x, r + t)$ is at distance $< t + a$ from some point in $B(x, r - a)$ and hence at distance $< t + 2a$ from some $z \in \Gamma(o) \cap B(x, r)$. Furthermore, if $z' \in \Gamma(o)$ is such that $d(x, z') \leq a$, then $N_o(z, t + 2a) = N_o(z', t + 2a) \leq N_o(x, t + 3a)$ for all $z \in \Gamma(o)$. Hence,

$$N_o(x, r + t) \leq N_o(x, r) N_o(x, t + 3a).$$

Now if $k \geq 0$ is the integer such that $kt < r \leq (k + 1)t$, then

$$N_o(x, r) \leq N_o(x, t + kt) \leq N_o(x, t) N_o(x, t + 3a)^k$$

and $\log(N_o(x, r)) \leq \log(N_o(x, t)) + \frac{r}{t} \log(N_o(x, t + 3a))$. Thus

$$\limsup_{r \rightarrow \infty} \frac{1}{r} \log(N_o(x, r)) \leq \frac{1}{t} \log(N_o(x, t + 3a))$$

for all $t > a$, and the existence of $\lim_{r \rightarrow \infty} \frac{1}{r} \log(N_o(x, r))$ follows. To complete the proof, one can check that there exist constants $b, C > 0$ such that

$$C^{-1} N_o(x, r - b) \leq \text{vol}(B(y, r)) \leq C N_o(x, r + b)$$

for all $y \in X$ and $r > b$. It suffices to consider the case that $d(x, o), d(y, o) \leq a$, and the two inequalities then follow by a ball packing or covering argument, respectively. $\square$

For example, if M has constant sectional curvature $\kappa \leq 0$, then

$$\text{vol}(B(x, r)) = V_{n, \kappa}(r) = \omega_{n-1} \int_0^r \text{sn}_\kappa^{n-1}(t) dt$$

(see Section 6.1) and hence $h(M) = (n - 1)\sqrt{-\kappa}$. Later we shall also compute the entropy of compact locally symmetric spaces of negative curvature. Some finer results on volume growth in the case of variable non-positive curvature, involving also polynomial terms, can be found in [Kn1997].

Lemma 8.14 (Poincaré series) *Let again $o, x \in X$, and let $s \geq 0$. The series*

$$P_{o,s}(x) := \sum_{\gamma \in \Gamma} e^{-s d(x, \gamma(o))}$$

associated with the group Γ of deck transformations of $\pi: X \rightarrow M$ converges if $s > h(M)$ and diverges if $s \leq h(M)$.

Proof: Let $r > 0$. Suppose that $0 \leq r_1 < r_2 < \dots < r_k$ are the values of $\gamma \mapsto d(x, \gamma(o))$ less than r . Put $r_{k+1} := r$, and let $n_i := N_o(x, r_i)$ be the number of orbit points in the open ball $B(x, r_i)$ for $i = 1, \dots, k+1$. There are $n_{i+1} - n_i$ points at distance r_i from x for $i = 1, \dots, k$, where $n_1 = 0$. Now

$$\sum_{i=1}^k (n_{i+1} - n_i) e^{-sr_i} = \sum_{i=1}^k n_{i+1} (e^{-sr_i} - e^{-sr_{i+1}}) + n_{k+1} e^{-sr_{k+1}},$$

and $n_{i+1} = N_o(x, t)$ for all $t \in (r_i, r_{i+1}]$. The corresponding partial sum of $P_{o,s}(x)$ can therefore be written as

$$\sum_{\gamma(o) \in B(x,r)} e^{-s d(x, \gamma(o))} = s \int_0^r N_o(x, t) e^{-st} dt + N_o(x, r) e^{-sr}.$$

If $s > h := h(M)$ and $\epsilon := \frac{1}{2}(s - h)$, then $\frac{1}{t} \log(N_o(x, t)) \leq h + \epsilon = s - \epsilon$ for sufficiently large t , so

$$N_o(x, t) e^{-st} \leq e^{-\epsilon t},$$

and the series converges. If $0 < s \leq h$, then by the proof of Lemma 8.13 there exists a constant $a > 0$ such that $h \leq \frac{1}{t-3a} \log(N_o(x, t))$ for $t > 4a$, thus

$$N_o(x, t) e^{-st} \geq N_o(x, t) e^{-ht} \geq e^{-3ah},$$

and the series diverges. If $s = 0$, then $P_{o,s}(x) = \#\Gamma = \infty$ by assumption. $\square$

The Poincaré series $P_{o,s}(x)$ is meaningful for any countable group Γ acting by isometries and with unbounded orbits on some metric space X . Notice that if y is another point in X , then

$$e^{-s d(x,y)} P_{o,s}(x) \leq P_{o,s}(y) \leq e^{s d(x,y)} P_{o,s}(x),$$

as $-d(x, y) - d(x, \gamma(o)) \leq -d(y, \gamma(o)) \leq d(x, y) - d(x, \gamma(o))$. The *critical exponent* of Γ is defined as the number

$$h := \sup\{s \geq 0 : P_{o,s}(x) = \infty\} \in [0, \infty],$$

which is independent of o and x . The group Γ is said to be of *divergence type* if $h < \infty$ and $P_{o,h}(x)$ diverges for all $x \in X$. Thus, the above results show that if X is the universal cover of a closed manifold M , then the group of deck transformations is of divergence type with critical exponent equal to the volume entropy of M .

We now assume again that X is a Hadamard manifold, or a proper CAT(0) space. Let $\Gamma \subset \text{Isom}(X)$ be any group of isometries. The *limit set* $\Lambda(\Gamma)$ of Γ consists of all points in X_∞ belonging to the closure, in $\bar{X}$, of the orbit $\Gamma(o)$. This is independent of the choice of base point o , and if Γ acts cocompactly, then $\Lambda(\Gamma) = X_\infty$. For a constant $\delta > 0$, a family $(\mu_x)_{x \in X}$ of non-trivial, finite Borel

measures on X_∞ is called a δ -dimensional conformal density for Γ if the measures are mutually absolutely continuous with Radon–Nikodym derivatives

$$\frac{d\mu_y}{d\mu_x}(\xi) = e^{\delta B_\xi(x,y)}$$

at $\xi \in X_\infty$ and the map $x \mapsto \mu_x$ is Γ -equivariant, that is,

$$\gamma_{\infty*}\mu_x = \mu_{\gamma(x)}$$

for all $\gamma \in \Gamma$ and $x \in X$.

Theorem 8.15 (conformal density) *Let X be a Hadamard manifold, or a proper CAT(0) space, and let $\Gamma \subset \text{Isom}(X)$ be a countable group with unbounded orbits. If Γ is of divergence type with critical exponent $h > 0$, then there exists an h -dimensional conformal density $(\mu_x)_{x \in X}$ for Γ such that $\mu_x(X_\infty \setminus \Lambda(\Gamma)) = 0$ for all $x \in X$.*

The following construction was devised by Patterson [Pat1976] for the hyperbolic plane and was then further developed by Sullivan [Su1979] for H^n . Various generalizations are discussed in [Ka1990], [Coo1993], [BurM1996], and [Al1999], among others. The assumption that Γ is of divergence type is not necessary for the theorem to hold, as long as $h \in (0, \infty)$, but simplifies the proof.

Proof: Fix a base point $o \in X$. For $x \in X$ and $s > h$, consider the measures

$$\mu_{x,s} := \frac{1}{P_{o,s}(o)} \sum_{\gamma \in \Gamma} e^{-s d(x, \gamma(o))} \delta_{\gamma(o)},$$

where $\delta_{\gamma(o)}$ denotes the Dirac measure at $\gamma(o)$. The total mass satisfies

$$\mu_{x,s}(X) = \frac{P_{o,s}(x)}{P_{o,s}(o)} \leq e^{s d(o,x)}.$$

Let first $x = o$. Since $\bar{X} = X \cup X_\infty$ is compact, it follows that for some decreasing sequence $s_i \rightarrow h$, the measures μ_{o,s_i} converge weakly to a probability measure μ_o on $\bar{X}$, concentrated on the closure of $\Gamma(o)$ in $\bar{X}$. In fact, because $P_{o,s_i}(o)$ tends monotonically to $P_{o,h}(o) = \infty$, we have $\mu_o(X) = 0$, and so μ_o must be concentrated on the limit set $\Lambda(\Gamma) \subset X_\infty$.

Next we show that for any other point $x \in X$, the measures μ_{x,s_i} , for the same sequence (s_i) , converge weakly to a measure μ_x such that $d\mu_x = e^{hB \cdot (o,x)} d\mu_o$. Given $x \in X$, define $f(y) := d(o, y) - d(x, y)$ for all $y \in X$; then

$$d\mu_{x,s} = e^{sf} d\mu_{o,s}$$

for $s > h$. By Proposition 8.4, f extends to a continuous function on $\bar{X}$ such that $f(\xi) = B_\xi(o, x)$ for all $\xi \in X_\infty$. Let $\phi \in C(\bar{X})$ and $\lambda > 1$. If i is sufficiently large, then $\frac{1}{\lambda} \leq e^{(s_i-h)f} \leq \lambda$ and hence

$$\frac{1}{\lambda} \int_{\bar{X}} \phi e^{hf} d\mu_{o, s_i} \leq \int_{\bar{X}} \phi d\mu_{x, s_i} \leq \lambda \int_{\bar{X}} \phi e^{hf} d\mu_{o, s_i}.$$

Since $\mu_{o, s_i} \rightarrow \mu_o$ weakly, and $\lambda > 1$ is arbitrary, the two integrals have the same limit $\int \phi e^{hf} d\mu_o$. Thus μ_{x, s_i} converges weakly to a measure μ_x satisfying

$$d\mu_x = e^{hf} d\mu_o = e^{hB_\cdot(o, x)} d\mu_o.$$

Using the same relation for another point $y \in X$, we conclude that

$$d\mu_y = e^{hB_\cdot(o, y)} d\mu_o = e^{hB_\cdot(x, y)} e^{hB_\cdot(o, x)} d\mu_o = e^{hB_\cdot(x, y)} d\mu_x,$$

as required.

It remains to show that if $\gamma \in \Gamma$, then $\gamma_{\infty*} \mu_x = \mu_{\gamma(x)}$. For all $\gamma' \in \Gamma$,

$$\gamma_*(e^{-s d(x, \gamma'(o))} \delta_{\gamma'(o)}) = e^{-s d(\gamma(x), \gamma\gamma'(o))} \delta_{\gamma\gamma'(o)};$$

therefore $\gamma_* \mu_{x, s_i} = \mu_{\gamma(x), s_i}$ for all i , and the result follows. $\square$

The Radon–Nikodym derivatives of a δ -dimensional conformal density agree, up to an extra exponent δ , with the conformal factors

$$\lim_{\xi' \rightarrow \xi} \frac{d_y(\xi, \xi')}{d_x(\xi, \xi')} = e^{B_\xi(x, y)}$$

of the visual metrics from Proposition 8.8. Indeed, for a proper CAT(−1) space X and a properly discontinuous and cocompact (or, more generally, convex cocompact) group $\Gamma \subset \text{Isom}(X)$ with exponential growth, the critical exponent $h > 0$ is equal to the Hausdorff dimension of the limit set $\Lambda(\Gamma)$ with respect to the metric d_x for any $x \in X$, and the family of the respective Hausdorff measures on $\Lambda(\Gamma)$ is an h -dimensional conformal density which agrees up to a constant factor with the density $(\mu_x)_{x \in X}$ from Theorem 8.15. See Théorème 2.7.4 in [Bou1995]. This shows in particular that the measures μ_x have no atoms.

8.5 The barycenter method

Theorem 8.15 assigns to every point $x \in X$ a non-trivial finite measure on X_∞ . We now discuss a reverse process for Hadamard manifolds of curvature at most -1 .

Theorem 8.16 (barycenter) *Let Y be a Hadamard manifold of curvature ≤ -1 , and let ω be a Borel probability measure on Y_∞ with the property that $\omega(\{\eta\}) < \frac{1}{2}$*

for all $\eta \in Y_\infty$. Fix a base point $p \in Y$ and write $b_\eta := B_\eta(\cdot, p)$ for the Busemann function of the ray from p to η . Then the function

$$\beta: Y \rightarrow \mathbb{R}, \quad \beta(y) := \int_{Y_\infty} b_\eta(y) d\omega(\eta),$$

is 1-Lipschitz, strictly convex, and has bounded sublevel sets, so it attains a global minimum at a uniquely determined point $\text{bar}(\omega) \in Y$.

Since $b_\eta(y) = B_\eta(y, p) = B_\eta(y, q) + B_\eta(q, p)$ for any $q \in Y$, it follows that the point $\text{bar}(\omega)$, the *barycenter* of ω , does not depend on the choice of the base point. Let γ be an isometry of Y , and consider the push-forward of ω under the induced homeomorphism $\gamma_\infty: Y_\infty \rightarrow Y_\infty$. By the change of variables formula,

$$\beta_\gamma(y) := \int_{Y_\infty} B_{\eta'}(y, p) d(\gamma_{\infty*}\omega)(\eta') = \int_{Y_\infty} B_{\gamma_\infty(\eta)}(y, p) d\omega(\eta),$$

and the second integrand equals $B_\eta(\gamma^{-1}(y), q)$ for $q := \gamma^{-1}(p)$. Thus $\beta_\gamma(y)$ is minimal when $\gamma^{-1}(y) = \text{bar}(\omega)$, and so

$$\text{bar}(\gamma_{\infty*}\omega) = \gamma(\text{bar}(\omega)).$$

That is, the barycenter map is $\text{Isom}(Y)$ -equivariant.

Proof of Theorem 8.16: Since every Busemann function b_η is 1-Lipschitz and convex, β has the same properties. To prove that β is in fact strictly convex, let $c: \mathbb{R} \rightarrow Y$ be a unit speed geodesic through $y := c(0)$. For $\eta \in Y_\infty$, let $\tau = \tau_\eta$ be the ray from y to η , and put $\alpha = \alpha_\eta := \angle_y(\tau'(0), c'(0))$. We first show that then

$$b_\tau(c(s)) \geq \log(\cosh(s) - \cos(\alpha) \sinh(s))$$

for all $s \in \mathbb{R}$. We use comparison with the hyperbolic upper half-plane (U^+, d) . Consider the geodesic $\bar{c}: \mathbb{R} \rightarrow U^+$ defined by $\bar{c}(s) := \tanh(a-s) + i(\cosh(a-s))^{-1}$, where $a \in \mathbb{R}$ is such that $\bar{y} := \bar{c}(0) = e^{i\alpha}$. Let $\bar{\tau}$ denote the upward vertical ray starting at $\bar{y}$, whose Busemann function satisfies $b_{\bar{\tau}}(u+iv) = -\log(v \cosh(a))$ for all $u+iv \in U^+$. Then $\angle_{\bar{y}}(\bar{\tau}'(0), \bar{c}'(0)) = \alpha$, and $d(c(s), \tau(t)) \geq d(\bar{c}(s), \bar{\tau}(t))$ for all $t \geq 0$ by hinge comparison, so $b_\tau(c(s)) \geq b_{\bar{\tau}}(\bar{c}(s))$ for all $s \in \mathbb{R}$. Furthermore,

$$b_{\bar{\tau}}(\bar{c}(s)) = \log\left(\frac{\cosh(a-s)}{\cosh(a)}\right) = \log(\cosh(s) - \tanh(a) \sinh(s)).$$

Since $\tanh(a) = \cos(\alpha)$, this gives the desired inequality. Now let $s > 0$. Then

$$\begin{aligned} b_\eta(c(s)) + b_\eta(c(-s)) - 2b_\eta(y) &= B_\eta(c(s), y) + B_\eta(c(-s), y) \\ &= b_\tau(c(s)) + b_\tau(c(-s)) \\ &\geq \log(\cosh^2(s) - \cos^2(\alpha) \sinh^2(s)) \\ &= \log(1 + \sin^2(\alpha) \sinh^2(s)). \end{aligned}$$

The set $A := Y_\infty \setminus \{c_{-\infty}, c_\infty\}$ has measure $\omega(A) > 0$ by the assumption on ω , and $\alpha_\eta \in (0, \pi)$ for all $\eta \in A$. It thus follows by integration that $\beta(c(s)) + \beta(c(-s)) - 2\beta(y) > 0$. This shows that β is strictly convex.

For every $r \geq 0$, the sublevel set $S_r := \{y \in Y : \beta(y) \leq r\}$ is convex and contains p , as $\beta(p) = 0$. We show that S_r is bounded. Let ϱ be a ray from p to a point $\xi \in Y_\infty$. If σ is the ray from p to a distinct point $\eta \in Y_\infty$, and if $s, t \geq 0$, then

$$d(\varrho(s), \sigma(t)) - t = s - 2(\varrho(s) | \sigma(t))_p \geq s - 2(\xi | \eta)_p$$

by monotonicity of the Gromov product, thus $b_\eta(\varrho(s)) \geq s + 2\log(\frac{1}{2}d_p(\xi, \eta))$ for the visual metric d_p . As $\omega(\{\xi\}) < \frac{1}{2}$, there exists an $\epsilon > 0$ such that the set $E := \{\eta : \frac{1}{2}d_p(\xi, \eta) \geq \epsilon\}$ has measure $\omega(E) > \frac{1}{2}$. Then $b_\eta(\varrho(s)) \geq s + 2\log(\epsilon)$ for $\eta \in E$ and $b_\eta(\varrho(s)) \geq -s$ for $\eta \in Y_\infty \setminus E$, hence

$$\beta(\varrho(s)) \geq (s + 2\log(\epsilon))\omega(E) - s(1 - \omega(E)) = 2\omega(E)(s + \log(\epsilon)) - s.$$

Since $2\omega(E) > 1$, this shows that β tends to infinity along ϱ . Thus every ray issuing from p leaves S_r , and since S_r is convex, an indirect argument shows that S_r must be bounded. $\square$

Using that Busemann functions in Hadamard manifolds are C^2 ([HeIH1977]), we can conclude from the above argument that

$$\text{Hess}(b_\eta)_y(v, v) = (b_\eta \circ c)''(0) \geq \sin^2(\alpha)$$

for the angle between the ray from y to $\eta \in Y_\infty$ and the initial velocity $v = c'(0) \in TY_y$ of the unit speed geodesic c (recall Example 2.11). Equality holds for $Y = H^n$. Notice that

$$\text{Hess}(\beta)_y(v, v) = \int_{Y_\infty} \text{Hess}(b_\eta)_y(v, v) d\omega(\eta),$$

thus it follows from the assumptions on ω that $\text{Hess}(\beta)_y$ is positive definite.

Theorem 8.11, Theorem 8.15 and Theorem 8.16 are the main constituents of the barycenter method used in the proof of the entropy rigidity theorem. To start the construction, we now make the following assumptions:

- X is an n -dimensional Hadamard manifold;
- $(\mu_x)_{x \in X}$ is a family of Borel measures on X_∞ without atoms such that $\mu_o(X_\infty) = 1$ for some base point $o \in X$ and

$$\frac{d\mu_x}{d\mu_o}(\xi) = e^{hB_\xi(o, x)}$$

for some fixed constant $h > 0$ and for all $x \in X$ and $\xi \in X_\infty$;

- Y is an n -dimensional Hadamard manifold of curvature ≤ -1 ;

- $f_\infty: X_\infty \rightarrow Y_\infty$ is a homeomorphism of the boundaries at infinity.

For the moment, it is irrelevant whether or not $(\mu_x)_{x \in X}$ is equivariant with respect to some group $\Gamma \subset \text{Isom}(X)$ and f_∞ is induced by a quasi-isometry $f: X \rightarrow Y$. For $x \in X$, we consider the probability measures

$$\nu_x := \frac{1}{\mu_x(X_\infty)} \mu_x, \quad \omega_x := f_{\infty*} \nu_x$$

on X_∞ and Y_∞ , respectively, and we associate with f_∞ the map

$$F: X \rightarrow Y, \quad F(x) := \text{bar}(\omega_x).$$

To describe F more explicitly, we fix again a base point $p \in Y$ and let $b_\eta := B_\eta(\cdot, p)$ denote the Busemann function of the ray from p to $\eta \in Y_\infty$. Likewise, we write $b_\xi := B_\xi(\cdot, o)$ for $\xi \in X_\infty$ and the base point $o \in X$. For $x \in X$, the function

$$\beta_x: Y \rightarrow \mathbb{R}, \quad \beta_x(y) := \int_{Y_\infty} b_\eta(y) d\omega_x(\eta),$$

is strictly convex and C^2 , and the image point $y = F(x)$ is uniquely characterized by the vector equation

$$0 = \text{grad } \beta_x(y) = \int_{Y_\infty} \text{grad } b_\eta(y) d\omega_x(\eta) \in TY_y.$$

Lemma 8.17 *The map $F: X \rightarrow Y$ is C^1 , and for every $x \in X$ and $y = F(x)$ the differential $dF_x: TX_x \rightarrow TY_y$ is characterized by the identity*

$$\text{Hess}(\beta_x)_y(dF_x(u), v) = h \int_{X_\infty} d(b_\xi)_x(u) d(b_{f_\infty(\xi)})_y(v) d\nu_x(\xi)$$

for all $(u, v) \in TX_x \times TY_y$.

Proof: Let first $x \in X$ and $y \in Y$ be arbitrary. By the change of variables formula for $\omega_x = f_{\infty*} \nu_x$, and since $\mu_x(X_\infty) d\nu_x = d\mu_x = e^{hB_\cdot(o, x)} d\mu_o = e^{-hb_\cdot(x)} d\mu_o$,

$$\mu_x(X_\infty) \text{grad } \beta_x(y) = \int_{X_\infty} e^{-hb_\xi(x)} \text{grad } b_{f_\infty(\xi)}(y) d\mu_o(\xi). \quad (1)$$

To show that F is C^1 , choose global normal coordinates around $y_0 = F(x_0)$, let $E_1, \dots, E_n$ denote the coordinate vector fields on Y , and put $e_i := E_i(y_0)$ and

$$G_i(x, y) := \mu_x(X_\infty) \langle \text{grad } \beta_x(y), E_i(y) \rangle.$$

By (1), as the Busemann functions are C^2 , the map $G := (G_1, \dots, G_n): X \times Y \rightarrow \mathbb{R}^n$ is C^1 , and

$$e_j(G_i(x_0, \cdot)) = \mu_{x_0}(X_\infty) \text{Hess}(\beta_{x_0})_{y_0}(e_j, e_i)$$

(recall Lemma 1.19). Since the Hessian is positive definite, and F is the unique map such that $G(x, F(x)) = 0$ for all $x \in X$, it follows from the implicit function theorem that F is C^1 .

Let now $x \in X$ and $y = F(x)$. Then the integral in (1) vanishes. Taking the derivative in direction of $u \in TX_x$, and reverting to the measure ν_x , we get the vector equation

$$\int_{X_\infty} D_{dF_x(u)} \operatorname{grad} b_{f_\infty(\xi)} d\nu_x(\xi) = h \int_{X_\infty} d(b_\xi)_x(u) \operatorname{grad} b_{f_\infty(\xi)}(y) d\nu_x(\xi).$$

Changing the variable back to $\eta \in Y_\infty$ on the left, and then taking the inner product with $v \in TY_y$, we get the claimed identity. $\square$

Our goal is to control the Jacobian $JF(x)$. To this end we introduce the self-adjoint endomorphisms H_x and K_x of $TY_y = TY_{F(x)}$ determined by the relations

$$\begin{aligned} \langle H_x(v), v \rangle &= \int_{Y_\infty} d(b_\eta)_y(v)^2 d\omega_x(\eta), \\ \langle K_x(v), v \rangle &= \operatorname{Hess}(\beta_x)_y(v, v) = \int_{Y_\infty} \operatorname{Hess}(b_\eta)_y(v, v) d\omega_x(\eta). \end{aligned}$$

Notice that K_x is invertible since $\operatorname{Hess}(\beta_x)_y$ is positive definite. Combining the identity of Lemma 8.17 with the Schwarz inequality we get the inequality

$$|\langle K_x \circ dF_x(u), v \rangle| \leq h \langle H_x(v), v \rangle^{1/2} \left(\int_{X_\infty} d(b_\xi)_x(u)^2 d\nu_x(\xi) \right)^{1/2} \quad (2)$$

for $y = F(x)$ and $(u, v) \in TX_x \times TY_y$.

Proposition 8.18 *For all $x \in X$,*

$$JF(x) \leq \frac{h^n}{n^{n/2}} \frac{\det(H_x)^{1/2}}{\det(K_x)}.$$

Furthermore, H_x has trace one.

Proof: We assume that dF_x is invertible; otherwise $JF(x) = 0$. Let (v_i) be an orthonormal basis of TY_y that diagonalizes H_x , and put $u'_i := (K_x \circ dF_x)^{-1}(v_i)$. Apply the Gram–Schmidt orthonormalization process to (u'_i) to obtain an orthonormal basis (u_j) of TX_x . Then every vector u_j is a linear combination $\sum_{i=1}^j \lambda_{ij} u'_i$, thus the matrix (λ_{ij}) of $K_x \circ dF_x$ with respect to (u_j) and (v_i) is triangular,

$$K_x \circ dF_x(u_j) = \sum_{i=1}^j \lambda_{ij} v_i.$$

The Jacobian $JF(x)$ may be computed as the absolute value of the determinant of the matrix of dF_x with respect to these bases. Hence

$$\det(K_x) JF(x) = |\det(\lambda_{ij})| = \prod_{i=1}^n |\langle K_x \circ dF_x(u_i), v_i \rangle|.$$

Putting $B_i := \int_{X_\infty} d(b_\xi)_x(u_i)^2 d\nu_x(\xi)$, we conclude from (2) and the choice of (v_i) that

$$\det(K_x) JF(x) \leq h^n \det(H_x)^{1/2} \prod_{i=1}^n B_i^{1/2}.$$

Furthermore, $\prod_i B_i \leq (\frac{1}{n} \sum_i B_i)^n$, and $\sum_i B_i = 1$ because

$$\sum_{i=1}^n \langle \text{grad } b_\xi(x), u_i \rangle^2 = |\text{grad } b_\xi(x)|^2 = 1$$

and ν_x is a probability measure. This gives the desired inequality, and it follows similarly as in this last step that H_x has trace $\sum_i \langle H_x(v_i), v_i \rangle = 1$. $\square$

8.6 Rigidity results

In the following, Y denotes an n -dimensional symmetric space of negative sectional curvature, normalized so that the maximum of the curvatures is -1 . By the classification of these spaces, Y is either real hyperbolic space H^n , complex hyperbolic space $\mathbb{C}H^m$ for $n = 2m$, quaternionic hyperbolic space $\mathbb{H}H^m$ for $n = 4m$, or the hyperbolic Cayley plane $\mathbb{O}H^2$ of real dimension $n = 16$. In the last three cases, the sectional curvature ranges over the interval $[-4, -1]$. We let $k \in \{1, 2, 4, 8\}$ denote the real dimension of the corresponding division algebra, so that $n = km$ (where $m = 2$ in the case of the octonions). We remark that Y is *two-point homogeneous*, that is, for any two pairs of points (y, y') and (z, z') with $d(y, y') = d(z, z')$ there exists an isometry of Y mapping y to z and y' to z' . Furthermore, the curvature tensor R of Y is parallel (this characterizes locally symmetric spaces).

Let $c: \mathbb{R} \rightarrow Y$ be a unit speed geodesic with $c(0) = y$ and $c'(0) = v$, and let E be a parallel unit normal vector field along c . If we choose $E(0)$ so that the sectional curvature $\langle E(0), R(E(0), v)v \rangle$ assumes the minimal or maximal value $\lambda \in \{-4, -1\}$, then $E(0)$ is a corresponding eigenvector of the (self-adjoint) operator $R(\cdot, v)v$, and since R is parallel, the relation

$$R(E(t), c'(t))c'(t) = \lambda E(t)$$

holds simultaneously for all $t \in \mathbb{R}$. It follows that, for this choice of E , Jacobi fields of the form $Y = fE$ exist and are determined by the equation $f'' + \lambda f = 0$ exactly as in space forms (see Example 3.8).

In order to compute the volume of balls in Y , we now use the fact that there exists a parallel orthonormal frame $(E_i)_{i=0}^{n-1}$ along c with $E_0 = c'$ such that

$$\sec(\text{span}\{c'(t), E_i(t)\}) = \begin{cases} -4 & \text{for } i = 1, \dots, k-1, \\ -1 & \text{for } i = k, \dots, n-1. \end{cases}$$

The Jacobi fields Y_i along c with $Y_i(0) = 0$ and $Y_i'(0) = E_i(0)$ are given by

$$Y_i(t) = d(\exp_y)_{tv}(tE_i(0)) = \begin{cases} \text{sn}_{-4}(t)E_i(t) = \sinh(t) \cosh(t)E_i(t), \\ \text{sn}_{-1}(t)E_i(t) = \sinh(t)E_i(t) \end{cases}$$

in the two cases, and therefore the Jacobian of $\exp_y$ at the point tv , for any $t > 0$, satisfies

$$J \exp_y(tv) = \frac{\sinh^{n-1}(t) \cosh^{k-1}(t)}{t^{n-1}}.$$

Hence, the volume of a ball $B(y, r)$ in Y is

$$\text{vol}(B(y, r)) = \omega_{n-1} \int_0^r \sinh^{n-1}(t) \cosh^{k-1}(t) dt,$$

and the volume entropy of any compact locally symmetric space N with universal cover Y is

$$h(N) = n + k - 2.$$

Next we need an explicit expression for the Hessian of a Busemann function $b: Y \rightarrow \mathbb{R}$. When $Y = H^n$, we already know that $\text{Hess}(b)_y(u, u) = \sin^2(\alpha)$ for the angle between $\text{grad } b(y)$ and any unit vector $u \in TY_y$, and since $db_y(u)^2 = \cos^2(\alpha)$, we have

$$\text{Hess}(b)_y(v, v) = |v|^2 - db_y(v)^2$$

for all $v \in TY_y$. Lemma 8.19 below extends this identity to the complex and quaternionic hyperbolic spaces. This relies on the fact that the complex or quaternionic structure on Y induces parallel $(1, 1)$ -tensor fields

$$J_1, \dots, J_{k-1}: \Gamma(TY) \rightarrow \Gamma(TY) \quad (k \in \{2, 4\})$$

satisfying $J_i^2 = -\text{id}$ and $\langle J_i(v), J_i(w) \rangle = \langle v, w \rangle$ for all $v, w \in TY_y$; moreover $v, J_1(v), \dots, J_{k-1}(v)$ are pairwise orthogonal, and $\sec(\text{span}\{v, J_i(v)\}) = -4$ (compare Example 4.9 for $\mathbb{C}P^m$, the compact dual of $\mathbb{C}H^m$). As discussed in [Ru2024], the Cayley hyperbolic plane possesses no analogous structure, and this causes some complications in the proof of the entropy rigidity theorem. For this reason, we omit the Cayley case ($k = 8$) from now on.

Lemma 8.19 *For any Busemann function $b: Y \rightarrow \mathbb{R}$ and $v \in TY_y$,*

$$\text{Hess}(b)_y(v, v) = |v|^2 - db_y(v)^2 + \sum_{i=1}^{k-1} db_y(J_i(v))^2.$$

Proof: Let $(v_0, \dots, v_{n-1})$ be an orthonormal basis of TY_y such that $v_0 = \text{grad } b(y)$ and $v_i = J_i(v_0)$ for $i = 1, \dots, k-1$, and let $c: \mathbb{R} \rightarrow Y$ be the geodesic with $c'(0) = v_0$. We compute $D_{v_i} \text{grad } b$ for $i \in \{1, \dots, n-1\}$. Let E_i be the parallel vector field along c with $E_i(0) = v_i$, and consider the Jacobi field Y_i defined by

$$Y_i(t) = e^{\kappa_i t} E_i(t),$$

where $\kappa_i = 2$ if $i \in \{1, \dots, k-1\}$ and $\kappa_i = 1$ if $i \in \{k, \dots, n-1\}$. Notice that this is the unique Jacobi field along c with $Y_i(0) = v_i$ and $\lim_{t \rightarrow -\infty} |Y_i(t)| = 0$. Hence, for any curve $\sigma: (-\epsilon, \epsilon) \rightarrow Y$ with $\sigma'(0) = v_i$, the integral curves $\gamma_s: \mathbb{R} \rightarrow Y$ of $\text{grad } b$ through $\sigma(s) = \gamma_s(0)$ form a geodesic variation with variation vector field Y_i along c . It follows that

$$D_{v_i} \text{grad } b = \frac{D}{ds} \Big|_{s=0} ((\text{grad } b) \circ \sigma) = \frac{D}{dt} \Big|_{t=0} Y_i = \kappa_i v_i.$$

(In fact, since $\text{grad } b$ is the unit normal vector field of the horospheres, $\kappa_1, \dots, \kappa_{n-1}$ are the principal curvatures of the horosphere through y .) Hence,

$$\text{Hess}(b)_y(v_i, v_i) = \langle D_{v_i} \text{grad } b, v_i \rangle = \kappa_i$$

for $i = 1, \dots, n-1$, whereas $\text{Hess}(b)_y(v_0, v_0) = 0$. One can now check that the identity in the lemma holds for all basis vectors $v_0, \dots, v_{n-1}$. $\square$

We will also need the following auxiliary result expressed purely in terms of linear algebra (compare Lemma B.4 in [BesCG1995]).

Lemma 8.20 *Let H be a symmetric positive definite $n \times n$ matrix with trace one, where $n \geq 3$. Then, for every positive real number $r \leq n-2$,*

$$\frac{\det(H)}{\det\left(\frac{1}{n-1}(I - H)\right)^{r+1}} \leq n^{nr},$$

and equality holds if and only if $H = \frac{1}{n}I$ (where I is the identity matrix).

Proof: Consider the simplex $\Sigma := \{(\lambda_1, \dots, \lambda_n) \in (0, 1)^n : \sum_i \lambda_i = 1\}$ of the possible n -tuples of eigenvalues of H . For $t \in (0, 1)$, put

$$\phi(t) := \log(t) - (r+1) \log(1-t).$$

We want to show that

$$\Phi(\lambda_1, \dots, \lambda_n) := \sum_{i=1}^n \phi(\lambda_i) \leq n \phi\left(\frac{1}{n}\right),$$

with equality if and only if $\lambda_1 = \dots = \lambda_n = \frac{1}{n}$.

Suppose that Φ attains a local maximum at $\lambda = (\lambda_1, \dots, \lambda_n)$. Then $\text{grad } \Phi(\lambda) = c(1, \dots, 1)$ for some $c \in \mathbb{R}$, thus $\phi'(\lambda_i) = c$ for $i = 1, \dots, n$. Furthermore, for every pair of distinct canonical basis vectors $e_i, e_j \in \mathbb{R}^n$,

$$\begin{aligned} 0 &\geq \frac{d^2}{(ds)^2} \Big|_{s=0} \Phi(\lambda + s(e_i - e_j)) = \frac{d}{ds} \Big|_{s=0} (\phi'(\lambda_i + s) - \phi'(\lambda_j - s)) \\ &= \phi''(\lambda_i) + \phi''(\lambda_j). \end{aligned}$$

One readily checks that there is a $t \in (0, 1)$ such that $\phi'' < 0$ on $(0, t)$ and $\phi'' > 0$ on $(t, 1)$, thus there are unique numbers $a \in (0, t]$ and $b \in [t, 1)$ with $\phi'(a) = c = \phi'(b)$. If $a < b$, then $\phi''(b) > 0$ and hence at most one eigenvalue is equal to b . Hence, in any case, there is an index l such that $\lambda_i = a$ for all $i \neq l$, whereas $\lambda_l \in \{a, b\}$. The equality

$$\frac{1 + ar}{a(1 - a)} = \phi'(a) = \phi'(\lambda_l) = \frac{1 + \lambda_l r}{\lambda_l(1 - \lambda_l)}$$

implies that $(\lambda_l - a)(1 - \lambda_l - a - a\lambda_l r) = 0$, and the second factor is positive since $1 - \lambda_l = (n - 1)a$ and $n - 2 \geq r > 0$. Thus, $\lambda_l = a$ as well, and $\lambda_i = \frac{1}{n}$ for all i .

It remains to analyze the values near the relative boundary of Σ . Notice that if $j \in \{1, \dots, n\}$ is fixed, then $\prod_{i \neq j} \lambda_i \leq \left(\frac{1}{n-1}(1 - \lambda_j)\right)^{n-1}$ by the arithmetic–geometric mean inequality, thus

$$\frac{\prod_i \lambda_i}{\prod_i \left(\frac{1}{n-1}(1 - \lambda_i)\right)^{r+1}} \leq (n-1)^{nr+1} \frac{\lambda_j(1 - \lambda_j)^{n-r-2}}{\prod_{i \neq j} (1 - \lambda_i)^{r+1}}.$$

For $\lambda_j \rightarrow 0$, this last term tends to zero, provided that all λ_i with $i \neq j$ remain bounded away from 1. On the other hand, if $\lambda_j \geq 1 - \epsilon$ for some $\epsilon \in (0, 1)$, then $\lambda_i \leq \epsilon$ for all $i \neq j$, and we get $(n-1)^{nr+1}(1 - \epsilon)^{(r+1)(1-n)}\epsilon^{n-r-2}$ as an upper bound. For $\epsilon \rightarrow 0$, this tends again to zero if $r < n - 2$ and to $(n-1)^{nr+1}$ if $r = n - 2$, and we must check that this number is smaller than n^{nr} . Indeed, since $n > 2$,

$$\left(\frac{n}{n-1}\right)^{nr} = \left(1 + \frac{1}{n-1}\right)^{n(n-2)} > 1 + \frac{n(n-2)}{n-1} > n-1.$$

This yields the result. □

We now return to the assumptions and the notation used in Proposition 8.18, with the addition that Y and $k \in \{1, 2, 4\}$ are as above, $n = \dim(X) = \dim(Y) \geq 3$, and the assumptions on $(\mu_x)_{x \in X}$ are satisfied for $h = n + k - 2$.

Proposition 8.21 *With these assumptions, $JF(x) \leq 1$ for all $x \in X$, and if $JF(x) = 1$ for some x , then the differential dF_x is a linear isometry.*

Proof: Let again $y = F(x)$. Integrating the identity from Lemma 8.19 for the Busemann function b_η against ω_x , we get the relation

$$K_x = I - H_x - \sum_{i=1}^{k-1} J_i H_x J_i,$$

where I is the identity map on TY_y . Putting $L_x := \frac{1}{n-1}(I - H_x)$ and recalling that $h = n + k - 2$, we write $\frac{1}{h}K_x$ as a convex combination

$$\frac{1}{h}K_x = \frac{n-1}{h}L_x + \sum_{i=1}^{k-1} \frac{1}{h}(-J_i H_x J_i)$$

(in case $k = 1$ this reads simply $\frac{1}{h}K_x = L_x$). We assume that $JF(x) > 0$. Then it follows from Proposition 8.18 that both H_x and L_x have positive eigenvalues and trace one. Now if A, B are two symmetric positive definite $n \times n$ matrices, then

$$\det((1-t)A + tB) \geq \det(A)^{1-t} \det(B)^t$$

for all $t \in [0, 1]$; that is, the determinant is log-concave on this set of matrices (see, for example, Lemma B.2 in [BesCG1995]). We conclude that

$$\det\left(\frac{1}{h}K_x\right) \geq \det(L_x)^{(n-1)/h} \det(H_x)^{(k-1)/h}.$$

In combination with Proposition 8.18, this gives the inequality

$$JF(x) \leq \frac{\det(H_x)^{1/2-(k-1)/h}}{n^{n/2} \det(L_x)^{(n-1)/h}}.$$

This expression turns out to be less than or equal to one by Lemma 8.20 for $r := \frac{h}{n-k}$, independently of x . The exponents of $\det(H_x)$ and $\det(L_x)$ are

$$\frac{1}{2} - \frac{k-1}{h} = \frac{1}{2r}, \quad \frac{n-1}{h} = \frac{r+1}{2r};$$

furthermore $r \geq 1$, with equality for $Y = H^n$, and $r+1 = \frac{2}{n-k}(n-1) \leq n-1$, with equality for H^3 and $\mathbb{C}H^2$.

Suppose now that $JF(x) = 1$. Then $H_x = \frac{1}{n}I$ by Lemma 8.20, hence $L_x = \frac{1}{n}I$, and $K_x = \frac{h}{n}I$. For $u \in TX_x$ and $v := dF_x(u)$, the inequality stated before Proposition 8.18 now reduces to

$$|dF_x(u)|^2 \leq n \int_{X_\infty} d(b_\xi)_x(u)^2 d\nu_x(\xi).$$

Let (u_i) be an orthonormal basis of TX_x , and let $G := (\langle dF_x(u_i), dF_x(u_j) \rangle)$ be the Gram matrix. Then it follows that $\text{trace}(G) \leq n$, again since $|\text{grad } b_\xi(x)| = 1$ and ν_x is a probability measure, and so

$$1 = JF(x)^2 = \det(G) \leq \left(\frac{1}{n} \text{trace}(G)\right)^n \leq 1.$$

The equality forces $G = (\delta_{ij})$, that is, dF_x is an isometry. $\square$

We are now ready for the proof of the main results.

Theorem 8.22 (Besson–Courtois–Gallot 1995) *Suppose that M and N are two closed negatively curved Riemannian manifolds of the same dimension $n \geq 3$, with isomorphic fundamental groups, and N is locally symmetric. Then*

$$h(M)^n \operatorname{vol}(M) \geq h(N)^n \operatorname{vol}(N),$$

and if $h(M) = h(N)$ and $\operatorname{vol}(M) = \operatorname{vol}(N)$, then M and N are isometric.

As above, we omit the case of the Cayley hyperbolic plane, which is discussed in [Ru2024].

Proof: The quantity $h(\cdot)^n \operatorname{vol}(\cdot)$ is scale invariant. We normalize the metrics so that the maximum of the sectional curvatures of N is -1 , thus $h(N) = n + k - 2$, and so that $h := h(M) = h(N)$. Let X and Y be the universal covers of M and N , respectively, and let $\iota: \Gamma_M \rightarrow \Gamma_N$ be an isomorphism of the groups of deck transformations. Fix base points $o \in X$ and $p \in Y$, and define the bijection $g: \Gamma_M(o) \rightarrow \Gamma_N(p)$ between the orbits by $g(\gamma(o)) = \iota(\gamma)(p)$ for all $\gamma \in \Gamma_M$. Extend g to a map $f: X \rightarrow Y$ by sending x to the unique circumcenter of the finite set $\{g(\gamma(o)) : d(x, \gamma(o)) = d(x, \Gamma_M(o))\}$ (recall Proposition 5.25). Notice that f satisfies

$$f \circ \gamma = \iota(\gamma) \circ f$$

for all $\gamma \in \Gamma_M$. It follows from Lemma 8.9 that f is a quasi-isometry.

By Theorem 8.11, f induces a homeomorphism $f_\infty: X_\infty \rightarrow Y_\infty$. Theorem 8.15, together with Lemma 8.14, shows that there exists an h -dimensional conformal density $(\mu_x)_{x \in X}$ for Γ_M . If $\gamma \in \Gamma_M$, then $\mu_x(X_\infty) = \mu_{\gamma(x)}(X_\infty)$, and hence the family of the normalized measures still satisfies $\gamma_{\infty*} \nu_x = \nu_{\gamma(x)}$ for all $x \in X$. These measures have no atoms (see the concluding remark in Section 8.4). Let now $F: X \rightarrow Y$ be the map defined by $F(x) := \operatorname{bar}(\omega_x)$ for $\omega_x := f_{\infty*} \nu_x$. Since

$$\omega_{\gamma(x)} = (f \circ \gamma)_{\infty*} \nu_x = (\iota(\gamma) \circ f)_{\infty*} \nu_x = \iota(\gamma)_{\infty*} \omega_x$$

and the barycenter construction is $\operatorname{Isom}(Y)$ -equivariant, it follows that

$$F \circ \gamma(x) = \operatorname{bar}(\omega_{\gamma(x)}) = \operatorname{bar}(\iota(\gamma)_{\infty*} \omega_x) = \iota(\gamma)(\operatorname{bar}(\omega_x)) = \iota(\gamma) \circ F(x)$$

for all $\gamma \in \Gamma_M$ and $x \in X$. As a consequence, F is at finite uniform distance from f , thus by continuity and Lemma 8.12 F is a surjective quasi-isometry, and F induces a surjective map $F_M: M \rightarrow N$. Finally, by Proposition 8.21,

$$\operatorname{vol}(N) \leq \int_N \#F_M^{-1}\{y\} d\operatorname{vol}(y) = \int_M JF_M(x) d\operatorname{vol}(x) \leq \operatorname{vol}(M),$$

and if $\operatorname{vol}(M) = \operatorname{vol}(N)$, then $JF_M(x) = 1$ for all $x \in M$, so $d(F_M)_x$ is a linear isometry, and $\#F_M^{-1}\{y\} = 1$ for all $y \in N$; hence F_M is an isometry. $\square$

In [BesCG1995] (and [Ru2024]) it is proved more generally that if $f: M \rightarrow N$ is a continuous map of degree $d > 0$ between two closed, oriented Riemannian n -manifolds, where N is locally symmetric with negative sectional curvature, then

$$h(M)^n \operatorname{vol}(M) \geq d \cdot h(N)^n \operatorname{vol}(N),$$

and for $n \geq 3$ equality implies that f is homotopic to a d -sheeted Riemannian covering, after a possible dilation of M . This result was conjectured by Gromov in the case when N has constant negative curvature (see p. 58 in [Gro1983]). Further generalizations to open manifolds of finite volume are obtained in [BoCS2005] and [St2006].

A first corollary of Theorem 8.22 is the rank one case of Mostow's rigidity theorem [Mos1973].

Theorem 8.23 (Mostow 1973) *Let M and N be two compact locally symmetric spaces of negative curvature and of the same dimension $n \geq 3$. If the fundamental groups are isomorphic and the metrics are normalized so that $\operatorname{vol}(M) = \operatorname{vol}(N)$, then M and N are isometric.*

Proof: By Theorem 8.22, $h(N) = h(M)$, and M and N are isometric (via F_M). $\square$

Mostow's original proof exploits the (quasi-)conformal geometry of the boundaries at infinity. We refer to [Bou2018] for an exposition of this strategy and a survey of further rigidity theorems obtained in this way.

Another corollary of Theorem 8.22 is the following result, also conjectured by Gromov (compare p. 13 in [Gro1982]).

Theorem 8.24 (Besson–Courtois–Gallot 1995) *Let (N, g) be a closed hyperbolic n -manifold, with constant sectional curvature -1 , and let $\tilde{g}$ be another Riemannian metric on N . If $\operatorname{Ric}_{\tilde{g}} \geq -(n-1)\tilde{g}$, then $\operatorname{vol}_{\tilde{g}}(N) \geq \operatorname{vol}_g(N)$.*

Proof: In case $n = 2$ this follows from the Gauss–Bonnet theorem. If $n \geq 3$, then

$$\operatorname{vol}_g(N) \leq \frac{h_{\tilde{g}}(N)^n}{h_g(N)^n} \operatorname{vol}_{\tilde{g}}(N) \leq \operatorname{vol}_{\tilde{g}}(N)$$

by Theorem 8.22 and Theorem 6.6 (Bishop). $\square$

One may ask further if the conclusion of Theorem 8.24 holds under the weaker assumption $\operatorname{scal}_{\tilde{g}} \geq -n(n-1)$; this is known as Schoen's conjecture (see p. 55 in [Gut2011], [Sa2022], and p. 3 in [Yu2023]).

Bibliography

- [Al1999] P. Albuquerque: Patterson–Sullivan theory in higher rank symmetric spaces. *Geom. Funct. Anal.* 9 (1999), no. 1, 1–28.
- [AleB1990] S. B. Alexander, R. L. Bishop: The Hadamard–Cartan theorem in locally convex metric spaces. *Enseign. Math.* 36 (1990), no. 3–4, 309–320.
- [Alex1951] 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.
- [Alex1955] A. D. Alexandrov: *Die innere Geometrie der konvexen Flächen.* Akademie-Verlag, Berlin 1955.
- [Ba1995] W. Ballmann: *Lectures on Spaces of Nonpositive Curvature.* DMV Seminar 25, Birkhäuser 1995.
- [Bas2018] G. Basso: Fixed point theorems for metric spaces with a conical geodesic bicombing. *Ergod. Th. Dynam. Sys.* 38 (2018), no. 5, 1642–1657.
- [BeN2007] I. D. Berg, I. G. Nikolaev: Quasilinearization and curvature of Aleksandrov spaces. *Geom. Dedicata* 133 (2008), 195–218.
- [Ber1955] 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.
- [Ber1962] M. Berger: An extension of Rauch’s metric comparison theorem and some applications. *Illinois J. Math.* 6 (1962), no. 4, 700–712.
- [BesCG1995] G. Besson, G. Courtois, S. Gallot: Entropies et rigidités des espaces localement symétriques de courbure strictement négative. *Geom. Funct. Anal.* 5 (1995), no. 5, 731–799.
- [BesCG1996] G. Besson, G. Courtois, S. Gallot: Minimal entropy and Mostow’s rigidity theorems. *Ergod. Th. Dynam. Syst.* 16 (1996), no. 4, 623–649.
- [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. L. Bishop, R. J. Crittenden: *Geometry of Manifolds.* Academic Press 1964.
- [Bl1916] W. Blaschke: *Kreis und Kugel.* Veit 1916. De Gruyter 1956.
- [BoCS2005] J. Boland, C. Connell, J. Souto: Volume rigidity for finite volume manifolds. *Amer. J. Math.* 127 (2005), no. 3, 535–550.
- [Bou1995] M. Bourdon: Structure conforme au bord et flot géodésique d’un $CAT(-1)$ -espace. *Enseign. Math.* 41 (1995), no. 1–2, 63–102.

- [Bou2018] M. Bourdon: Mostow type rigidity theorems. In: L. Ji, A. Papadopoulos, S.-T. Yau (Eds.), *Handbook of Group Actions IV*, Adv. Lect. Math. (ALM) 41, Internat. Press 2018, pp. 139–188.
- [BrH1999] M. R. Bridson, A. Haefliger: *Metric Spaces of Non-Positive Curvature*. Grundlehren der Math. Wissenschaften 319, Springer 1999.
- [BruNS2025] E. Bruè, A. Naber, D. Semola: Fundamental groups and the Milnor conjecture. *Ann. of Math.* 201 (2025), no. 1, 225–289.
- [BruT1972] F. Bruhat, J. Tits: Groupes réductifs sur un corps local. Chapitre premier. Données radicielles valuées. *Publications Math. IHES* 41 (1972), 5–251.
- [BuBI2001] D. Burago, Y. Burago, S. Ivanov: *A Course in Metric Geometry*. Graduate Studies in Math. 33, Amer. Math. Soc. 2001.
- [BuGP1992] Y. D. Burago, M. Gromov, G. Perelman: A. D. Alexandrov spaces with curvature bounded below. *Russian Math. Surveys* 47 (1992), no. 2, 1–58.
- [BurM1996] M. Burger, S. Mozes: CAT(−1)-spaces, divergence groups and their commensurators. *J. Amer. Math. Soc.* 9 (1996), no. 1, 57–93.
- [Bus1948] H. Busemann: Spaces with non-positive curvature. *Acta Math.* 80 (1948), 259–310.
- [Buse1985] P. Buser: A geometric proof of Bieberbach’s theorems on crystallographic groups. *Enseign. Math.* 31 (1985), no. 1-2, 137–145.
- [BuyS2007] S. Buyalo, V. Schroeder: *Elements of Asymptotic Geometry*. EMS Monographs in Math., Europ. Math. Soc. 2007.
- [By1970] W. P. Byers: On a theorem of Preissmann. *Proc. Amer. Math. Soc.* 24 (1970), 50–51.
- [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 non-negative 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.
- [Coo1993] M. Coornaert: Mesures de Patterson-Sullivan sur le bord d’un espace hyperbolique au sens de Gromov. *Pacific J. Math.* 159 (1993), no. 2, 241–270.
- [DeL2016] D. Descombes, U. Lang: Flats in spaces with convex geodesic bicomings. *Anal. Geom. Metric Spaces* 4 (2016), 68–84.
- [DoC1992] M. P. do Carmo: *Riemannian Geometry*. Birkhäuser 1992, 1993.
- [EbON1973] P. Eberlein, B. O’Neill: Visibility manifolds. *Pacific J. Math.* 46 (1973), 45–109.
- [FeN2003] W. Fenchel, J. Nielsen: *Discontinuous Groups of Isometries in the Hyperbolic Plane*. Edited by Asmus L. Schmidt. De Gruyter Studies in Math. 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.
- [GroW1971] D. Gromoll, J. A. Wolf: Some relations between the metric structure and the algebraic structure of the fundamental group in manifolds of nonpositive curvature. *Bull. Amer. Math. Soc.* 77 (1971), 545–552.
- [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.
- [Gro1981c] 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.
- [Gro1982] M. Gromov: Volume and bounded cohomology. *Publ. Math. Inst. Hautes Études Sci.* 56 (1982), 5–99.
- [Gro1983] M. Gromov: Filling Riemannian manifolds. *J. Differential Geom.* 18 (1983), no. 1, 1–147.
- [Gro1987] M. Gromov: Hyperbolic groups. In: S. M. Gersten (Ed.), *Essays in Group Theory*, MSRI Publ. 8, Springer 1987, pp. 75–263.
- [Gro1987] 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.
- [Gui2000] L. Guijarro: On the metric structure of open manifolds with nonnegative curvature. *Pacific J. Math.* 196 (2000), no. 2, 429–444.
- [GuiW2000] L. Guijarro, G. Walschap: The metric projection onto the soul. *Trans. Amer. Math. Soc.* 352 (2000), no. 1, 55–69.
- [Gut2011] L. Guth: Volumes of balls in large Riemannian manifolds. *Ann. of Math.* 173 (2011), no. 1, 51–76.
- [Ha1990] U. Hamenstädt: Entropy-rigidity of locally symmetric spaces of negative curvature. *Ann. of Math.* 131 (1990), no. 1, 35–51.
- [HeIH1977] E. Heintze, H.-C. Im Hof: Geometry of horospheres. *J. Differential Geometry* 12 (1977), no. 4, 481–491.
- [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.
- [JoN1935] P. Jordan, J. von Neumann: On inner products in linear metric spaces. *Ann. Math.* 36 (1935), 719–723.
- [Jos1995] J. Jost: *Riemannian Geometry and Geometric Analysis*. Springer 1995.

- [Jos1997] J. Jost: Nonpositive Curvature: Geometric and Analytic Aspects. Lectures in Math. ETH Zürich, Birkhäuser 1997.
- [Ka1990] V. A. Kaimanovich: Invariant measures of the geodesic flow and measures at infinity on negatively curved manifolds. *Ann. Inst. H. Poincaré Sect. A* 53 (1990), no. 4, 361–393.
- [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.
- [Kn1997] G. Knieper: On the asymptotic geometry of nonpositively curved manifolds. *Geom. Funct. Anal.* 7 (1997), no. 4, 755–782.
- [KoS1993] N. J. Korevaar, R. M. Schoen: Sobolev spaces and harmonic maps for metric space targets. *Comm. Anal. Geom.* 1 (1993), no. 3-4, 561–659.
- [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.
- [LaS1997] U. Lang, V. Schroeder: Quasiflats in Hadamard spaces. *Ann. Sci. École Norm. Sup.* 30 (1997), no. 3, 339–352.
- [LaS2013] U. Lang, V. Schroeder: On Toponogov's comparison theorem for Alexandrov spaces. *Enseign. Math.* 59 (2013), no. 3–4, 325–336.
- [LawY1972] H. B. Lawson, S. T. Yau: Compact manifolds of nonpositive curvature. *J. Differential Geom.* 7 (1972), 211–228.
- [Liu2013] G. Liu: 3-manifolds with nonnegative Ricci curvature. *Invent. Math.* 193 (2013), no. 2, 367–375.
- [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.
- [Ma1979] A. Manning: Topological entropy for geodesic flows. *Ann. of Math.* 110 (1979), no. 3, 567–573.
- [Mi1968] J. Milnor: A note on curvature and fundamental group. *J. Differential Geometry* 2 (1968), 1–7.
- [Mo1924] H. M. Morse: A fundamental class of geodesics on any closed surface of genus greater than one. *Trans. Amer. Math. Soc.* 26 (1924), no. 1, 25–60.
- [Mos1968] 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.
- [Mos1973] 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.

- [Pa2020] J. Pan: A proof of Milnor conjecture in dimension 3. *J. Reine Angew. Math.* 758 (2020), 253–260.
- [Pap2005] A. Papadopoulos: Metric Spaces, Convexity and Nonpositive Curvature. IRMA Lectures in Math. and Theor. Physics 6, Europ. Math. Soc. 2005.
- [Pat1976] S. J. Patterson: The limit set of a Fuchsian group. *Acta Math.* 136 (1976), no. 3–4, 241–273.
- [Pe1994] G. Perelman: Proof of the soul conjecture of Cheeger and Gromoll. *J. Differential Geom.* 40 (1994), no. 1, 209–212.
- [Pet1984] S. Peters: Cheeger’s finiteness theorem for diffeomorphism classes of Riemannian manifolds. *J. Reine Angew. Math.* 349 (1984), 77–82.
- [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.
- [Re1968] Y. G. Reshetnyak: Inextensible mappings in a space of curvature no greater than K . *Siberian Math. J.* 9 (1968), no. 4, 683–689.
- [Ru2024] Y. Ruan: The Cayley hyperbolic space and volume entropy rigidity. *Math. Z.* 306 (2024), no. 1, article no. 4, 30 pp.
- [Sa2022] S. Sabourau: Macroscopic scalar curvature and local collapsing. *Ann. Sci. Éc. Norm. Supér.* 55 (2022), no. 4, 919–936.
- [Sat2009] T. Sato: An alternative proof of Berg and Nikolaev’s characterization of CAT(0)-spaces via quadrilateral inequality. *Arch. Math.* 93 (2009), no. 5, 487–490.
- [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.
- [St2006] P. A. Storm: The minimal entropy conjecture for nonuniform rank one lattices. *Geom. Funct. Anal.* 16 (2006), no. 4, 959–980.
- [Su1979] D. Sullivan: The density at infinity of a discrete group of hyperbolic motions. *Inst. Hautes Études Sci. Publ. Math.* 50 (1979), 171–202.
- [Sv1955] A. S. Švarc: A volume invariant of coverings (Russian). *Dokl. Ak. Nauk SSSR* 105 (1955), 32–34.
- [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.
- [Wi2007] B. Wilking: A duality theorem for Riemannian foliations in nonnegative sectional curvature, Geom. Funct. Anal. 17 (2007), no. 4, 1297–1320.
- [Wo2011] J. A. Wolf: Spaces of Constant Curvature. Sixth Edition. AMS Chelsea Publishing 2011.
- [Yi1988] J.-W. Yim: Distance nonincreasing retraction on a complete open manifold of nonnegative sectional curvature. Ann. Global Anal. Geom. 6 (1988), no. 2, 191–206.
- [Yu2023] W. Yuan: Volume comparison with respect to scalar curvature. Anal. PDE 16 (2023), no. 1, 1–34.