Contents

Comparison Theorems in Riemannian Geometry

J.-H. Eschenburg

0. Introduction

The subject of these lecture notes is comparison theory in Riemannian geometry: What can be said about a complete Riemannian manifold when (mainly lower) bounds for the sectional or Ricci curvature are given? Starting from the comparison theory for the Riccati ODE which describes the evolution of the principal curvatures of equidistant hypersurfaces, we discuss the global estimates for volume and length given by Bishop-Gromov and Toponogov. An application is Gromov's estimate of the number of generators of the fundamental group and the Betti numbers when lower curvature bounds are given. Using convexity arguments, we prove the "soul theorem" of Cheeger and Gromoll and the sphere theorem of Berger and Klingenberg for nonnegative curvature. If lower Ricci curvature bounds are given we exploit subharmonicity instead of convexity and show the rigidity theorems of Myers-Cheng and the splitting theorem of Cheeger and Gromoll. The Bishop-Gromov inequality shows polynomial growth of finitely generated subgroups of the fundamental group of a space with nonnegative Ricci curvature (Milnor). We also discuss briefly Bochner's method.

The leading principle of the whole exposition is the use of convexity methods. Five ideas make these methods work: The comparison theory for the Riccati ODE, which probably goes back to L.Green [15] and which was used more systematically by Gromov [20], the triangle inequality for the Riemannian distance, the method of support function by Greene and Wu [16],[17],[34], the maximum principle of E.Hopf, generalized by E.Calabi [23], [4], and the idea of critical points of the distance function which was first used by Grove and Shiohama [21]. We have tried to present the ideas completely without being too technical.

These notes are based on a course which I gave at the University of Trento in March 1994. It is a pleasure to thank Elisabetta Ossanna and Stefano Bonaccorsi who have worked out and typed part of these lectures. We also thank Evi Samiou and Robert Bock for many valuable corrections.

Augsburg, September 1994

J.-H. Eschenburg

1. Covariant derivative and curvature.

Notation: By M we always denote a smooth manifold of dimension n. For $p \in M$, the tangent space at p is denoted by T_pM, and TM denotes the tangent bundle. If M' is another manifold and $f : M \to M'$ a smooth (i.e. C^∞) map, its differential at some point $p \in M$ is always denoted by $df_p : T_pM \to T_{f(p)}M'$. For $v \in T_pM$ we write $df_p(v) = df_p.v = \partial_v f$. If $c : I \to M$ is a (smooth) curve, we denote its tangent vector by $c'(t) = dc(t)/dt = dc_t.1 \in T_{c(t)}M$ (where $1 \in T_tI = \mathbb{R}$). If $f : M \to \mathbb{R}$, then $df_p \in (T_pM)^*$. If M is a Riemannian manifold, i.e. there exists a scalar product $< , >$ on any tangent space of M, this gives an isomorphism between T_pM and $(T_pM)^*$; the vector $\nabla f(p)$ corresponding to df_p is called the *gradient* of f.

Let M be a Riemannian manifold. We denote by $< , >$ the scalar product on M and we define the norm of a vector by

$$\|v\| = \sqrt{< v, v >},$$

the *length* of a curve $c : I \to M$ by

$$L(c) = \int_I \|c'(t)\| dt,$$

and the *distance* between $x, y \in M$ by

$$|x, y| = \inf \{L(c) \ ; \quad c : x \to y\} .$$

where $c : x \to y$ means that $c : [a, b] \to M$ with $c(a) = x$ and $c(b) = y$. If $L(c) = |x, y|$ for some $c : x \to y$, then c is called *shortest*. The open and closed metric balls are denoted by $B_r(p)$ and $D_r(p)$, i.e.

$$B_r(p) = \{x \in M; \ |x, p| < r\}, \quad D_r(p) = \{x \in M; \ |x, p| \le r\}.$$

Similarly, we define $B_r(A)$ for any closed subset $A \subset M$.

We denote by $\mathbf{X}(M)$ the set of vector fields on M.

Definition 1.1 The *Levi-Civita covariant derivative*

$$D : \mathbf{X}(M) \times \mathbf{X}(M) \to \mathbf{X}(M)$$

$$(X, Y) \to D_X Y,$$

is determined by the following properties holding for all functions $f, g \in C^\infty(M)$ and for all vector fields $X, X', Y, Y' \in \mathbf{X}(M)$:

1. $D_{(fX+gX')}Y = fD_XY + gD_{X'}Y$;
2. $D_X(fY + gY') = (\partial_X f)Y + fD_XY + (\partial_X g)Y' + gD_XY'$;
3. $D_XY - D_YX = [X, Y] =$ "Lie bracket";
4. $\partial_Z < X, Y >=< D_ZX, Y > + < X, D_ZY >$.

Definition 1.2 The *Riemannian curvature tensor* $(X, Y, Z) \mapsto R(X, Y)Z$ is defined as follows:

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

It satisfies certain algebraic identities ("*curvature identities*"), namely

$$< R(X, Y)Z, W >= - < R(Y, X)Z, W >= - < R(X, Y)W, Z >=< R(Z, W)X, Y >$$

and the Bianchi identity

$$R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0$$

(cf. [29]). In particular,

$$R_V := R(., V)V$$

is a self adjoint endomorphism of TM for any vector field V on M. Several notions of *curvature* are derived from this tensor:

1. *Sectional curvature* $K(\ ,\ )$: For every linearly independent pair of vectors $X, Y \in T_pM$,

$$K(X, Y) = \frac{< R(X, Y)Y, X >}{\|X\|^2\|Y\|^2 - < X, Y >^2}.$$

   K is defined on the space of two dimensional linear subspaces of T_pM (depending only on $span(X, Y)$).

2. *Ricci curvature*

$$Ric(X, Y) = \mathrm{trace}(Z \mapsto R(Z, X)Y).$$

   By the curvature identities, $Ric(X, Y) = Ric(Y, X)$. Ricci curvature in direction X is given by

$$Ric(X) = Ric(X, X)$$

   where X is a unit vector.

3. *Scalar curvature*

$$s = \mathrm{trace}(Ric) = \sum Ric(E_i, E_i)$$

where $\{E_i\}_{i=1}^n$ is a local orthonormal basis.

There is a close relationship between $R_V = R(., V)V$ and the sectional curvature: Let $\|V\| = 1$. For X orthogonal to V we have

$$< R_V X, X > = < R(X, V)V, X > = K(V, X)\|X\|^2$$

Hence the highest ("λ_+") and lowest ("λ_-") eigenvalues of R_V give a bound to $K(V, X)$, since

$$\lambda_-(R_V) \leq \frac{< R_V X, X >}{< X, X >} \leq \lambda_+(R_V).$$

Moreover, trace$(R_V) = Ric(V, V)$.

Let us come back to the covariant derivative. It is easy to see that for any $p \in M$, $(D_X Y)_p$ depends only on $dY_p.X(p)$ where the vector field Y is considered as a smooth map $Y : M \to TM$. Therefore, the covariant derivative is also defined if the vector fields X and Y are only partially defined. E.g. if $\gamma : I \to M$ is a smooth regular curve and Y is a *vector field along* γ, i.e. a smooth map $Y : I \to TM$ with

$$Y(t) \in T_{\gamma(t)}M$$

for all $t \in I$ (e.g. γ' is such a vector field), then

$$Y'(t) := \frac{DY(t)}{dt} := D_{\gamma'(t)}Y$$

is defined (just extend γ' and Y arbitrarly outside γ). Similar, if $\gamma : I_1 \times ... \times I_k \to M$ depends on k variables, we have k partial derivatives $\frac{\partial \gamma}{\partial t_i}$ and corresponding covariant derivatives $\frac{D}{\partial t_i}$ $(i = 1, ..., k)$ along γ. (Formally, a vector field along γ is a section of the pull-back bundle $\gamma^* TM$, and D induces a covariant derivative on this bundle.)

Definition 1.3 A vector field Y along a curve $\gamma : I \to M$ is called *parallel* if $Y' = 0$. A curve γ is a called a *geodesic* in M if γ' is parallel, i.e. if

$$(\gamma')' = D_{\gamma'}\gamma' = 0. \tag{1.1}$$

(1.1) is a 2^{nd} order ODE. In fact, if $x = (x^1, ..., x^n) : M \to \mathbb{R}^n$ is a coordinate chart with $E_i = \frac{\partial}{\partial x^i}$ and if we put

$$D_{E_i} E_j = \Gamma_{ij}^k E_k$$

(summation convention!), then $\gamma' = (\gamma^i)' E_i$ where $\gamma^i := x^i \circ \gamma$, and

$$\frac{D}{dt}\gamma' = (\gamma^i)'' E_i + (\gamma^j)' D_{\gamma'} E_j$$

with

$$D_{\gamma'} E_j = \gamma^k D_{E_k} E_j = (\gamma^k)' \Gamma^i_{kj} E_i,$$

hence (1.1) is equivalent to

$$(\gamma^i)'' + (\gamma^j)'(\gamma^k)'\Gamma^i_{kj} = 0$$

To some extent, Riemannian geometry is the theory of this ODE.

Definition 1.4 For any $v \in TM$ let γ_v denote the unique geodesic with $\gamma'(0) = v$. For $s, t \in \mathbb{R}$ with $|s|$ and $|t|$ small we have $\gamma_{sv}(t) = \gamma_v(st)$ by uniqueness for ODE's. Thus for $v \in TM$ with $\|v\|$ small enough,

$$\exp(v) := \gamma_v(1)$$

is defined and gives a smooth map $\exp : (TM)_0 \to M$ where $(TM)_0$ is a neighborhood of the zero section of TM. This is called the *exponential map* of M. M is called *(geodesically) complete* if exp is defined on all of TM. Fixing $p \in M$, we put $\exp_p = \exp|T_pM$.

Remark 1.5 The map $\exp_p$ is a diffeomorphism near the origin (in fact, $d(\exp_p)_0$ is the identity on T_pM), and it maps all the lines through the origin of T_pM onto the geodesics through the point $p \in M$. Thus, $\exp_p : T_pM \to M$ can serve a a coordinate map near p ("*exponential coordinates*") which preserves the covariant derivative at p, i.e. covariant differentiation at p is the same as taking the ordinary derivative at $0 \in T_pM$ in exponential coordinates. To see this, identify M and T_pM near p via $\exp_p$ and consider the ordinary derivative $D^o = \partial$ on T_pM. It satisfies the rules (1.),(2.) and (3.) of the Levi-Civita derivative (but not 4. for the Riemannian metric). Hence the difference $\Gamma = D - \partial$ is a tensor field, i.e. $\Gamma_X fY = f\Gamma_X Y$ for all functions f (by 2.), further it is symmetric, i.e. $\Gamma_X Y = \Gamma_Y X$ (by 3.), and it satisfies $\Gamma_v v = 0$ for all $v \in T_pM$, since D and ∂ have the same geodesics at p. Thus $\Gamma|_p = 0$.

2. Jacobi and Riccati equations; equidistant hypersurfaces.

Equation (1.1) is a nonlinear ODE which in general cannot be solved explicitly. Therefore, we consider its linearization. This is the ODE satisfied by a variation of solutions of (1.1), i.e. of geodesics. So let $\gamma(s,t) = \gamma_s(t)$ be a smooth one-parameter family of geodesics γ_s. Put $V = \frac{\partial\gamma}{\partial t} \in \mathbf{X}(\gamma_s)$ and $J = \frac{\partial\gamma}{\partial s}$. Then J is the variation vector field and V the tangent field of the geodesics γ_s, hence $D_V V = O$.

Fig. 1.

Then we have

$$J'' = \frac{D}{\partial t}\frac{D}{\partial t}J = \frac{D}{\partial t}\frac{D}{\partial t}\frac{\partial\gamma}{\partial s}.$$

We can interchange the order of differentiation, getting

$$J'' = \frac{D}{\partial s}\frac{D}{\partial t}\frac{\partial\gamma}{\partial t} + R(V,J)V,$$

$$J'' + R(J,V)V = 0. \tag{2.1}$$

Equation (2.1) is called *Jacobi equation*.

Definition 2.1 A vector field J along a geodesic γ is called a *Jacobi field* if it satisfies the Jacobi equation.

Remark 2.2 J is a Jacobi field along γ if and only if

$$J(t) = \left.\frac{d}{ds}\right|_0 \gamma_s(t) \tag{2.2}$$

for some one-parameter family of geodesics γ_s with $\gamma_0 = \gamma$.

Implication "$\Leftarrow$" was shown above. To prove the opposite implication, we have to construct the family γ_s. Let $\alpha(s) = \exp_{\gamma(0)} sJ(0)$. Let X be a vector field along α such that $X(0) = \gamma'(0)$ and $X'(0) = J'(0)$ and put

$$\gamma_s(t) = \exp_{\alpha(s)} tX(s) \qquad (2.3).$$

If we put

$$\tilde{J} = \frac{\partial}{\partial s}\Big|_0 \gamma_s,$$

then, by "$\Leftarrow$", $\tilde{J}$ satisfies the Jacobi equation. Since $\tilde{J}(0) = J(0)$ and

$$\tilde{J}'(0) = \frac{D}{\partial t}\frac{\partial}{\partial s}\gamma|_{(0,0)} = \frac{D}{\partial s}\frac{\partial}{\partial t}\gamma|_{(0,0)} = \frac{D}{\partial s}X(s)|_0 = X'(0) = J'(0),$$

we get $J = \tilde{J}$ by uniqueness of the solution.

Next, we want to split this 2^{nd} degree equation in a system of 1^{st} degree equations. To do this, we embed the 1-parameter family of geodesics describing the Jacobi field into an $(n-1)$-parameter family. I.e. we choose a suitable smooth map

$$\gamma : S \times I \to M$$

where S is an $(n-1)$-dimensional manifold, such that $\gamma_s(t) = \gamma(s,t)$ is a geodesic for any $s \in S$. If γ is a regular map, then $V = d\gamma(\frac{\partial}{\partial t})$ can be viewed as a vector field on an open subset of M with $D_V V = 0$, and the Jacobi fields J arising from variations in S-directions commute with V, i.e. we have $[J,V] = 0$ or

$$D_V J = A \cdot J \qquad (2.4)$$

where $A = DV$, i.e. $A \cdot X = D_X V$. This is the first of our equations: Knowing A, we recieve J by solving a 1^{st} order equation.

It remains to derive an equation for A. Let us consider first an arbitrary vector field V on M and let $A = DV$ as before. In general, the covariant derivative of a tensor field A is defined by

$$(D_V A)X = D_V(AX) - A \cdot D_V(X).$$

Hence we have

$$\begin{aligned}(D_V A)X &= D_V D_X V - A(D_X V + [V,X]) \\ &= D_X D_V V + R(V,X)V + D_{[V,X]}V - A^2 X - A[V,X] \\ &= D_X D_V V + R_V(X) - A^2 X \\ &.\end{aligned}$$

Therefore

$$D_V A + A^2 + R_V = D(D_V V). \tag{2.5}$$

If we suppose $D_V V = 0$ (i.e. the integral curves γ_s are geodesic), then we get an ODE for A, the so called *Riccati equation*

$$A' + A^2 + R_V = 0.$$

Thus we have split the Jacobi equation $J'' = -R_V J$ in two equations as follows:

$$J' = AJ \tag{2.6}$$

$$A' + A^2 + R_V = 0. \tag{2.7}$$

We note that the second equation can be solved independently of the first.

Let us consider now the important special case where $(DV)^* = DV$, that is

$$< D_X V, Y >=< X, D_Y V >$$

for all vector fields X,Y. Then V is locally a gradient, i.e. locally $V = \nabla f$ for some function $f : M \to \mathbb{R}$. Consequently, $< V, V >$ is constant, since

$$\partial_X < V, V >= 2 < D_X V, V >= 2 < D_V V, X >= 0.$$

Thus we may assume that $< V, V >= 1$. Now let us consider the level hypersurfaces

$$S_t = \{x \in M : \ f(x) = t\}.$$

Since $V = \nabla f \neq 0$, the S_t are regular hypersurfaces and $V|_{S_t}$ is the unit normal vector field on S_t. Thus in this case, our $(n-1)$-parameter family of geodesics $\gamma : S \times I \to M$ is given by

$$\gamma(s,t) = \exp(t - t_0)V(s)$$

where $S = S_{t_0}$ for some $t_0 \in I$, and $f(\gamma(s,t)) = t$, or in other words, $S_t = \phi_t(S)$ where $\phi_t(s) := \gamma(s,t)$. Such a family of hypersurfaces S_t is called *equidistant*, and the function $f - t_0$ is called the *signed distance function* of the hypersurface $S = S_{t_0}$. In fact we have

$$|f(x) - t_0| = |x, S| := \inf_{s \in S} |x, s| \tag{2.8}$$

for x in a small neighborhood of S. Namely, if $c : [a,b] \to \gamma(S \times I) \subset M$ is a curve with $c(a) \in S_{t_0}$ and $c(b) \in S_{t_1}$, then we have $c(u) = \gamma(s(u), t(u))$ with $t(a) = t_0$, $t(b) = t_1$, and

$$\|c'(u)\|^2 = \|d\gamma.s'(u)\|^2 + t'(u)^2 \geq t'(u)^2,$$

hence its length is

$$L(c) \geq \int_a^b |t'(u)| du \geq |t(b) - t(a)| \geq |t_0 - t_1|.$$

Fig. 2.

In this case, all the quantities discussed above have geometric meanings. The Jacobi fields $J(t) = d\gamma_{(s,t)}(x,0) = d\phi_t x$ for $x \in T_s S$ measure the change of the metric of $S_t = \phi_t(S)$ when t is changed; in fact, $\|J(t)\|/\|J(t_0)\|$ is the length distortion between S_t and S. Moreover, $A = DV$, restricted to the hypersurface S_t, is the *shape operator* of S_t since $V|S_t$ is a unit normal vector field on S_t. Its eigenvalues are called *principal curvatures*, their average the *mean curvature* of S_t. Since Equation (2.7) is nonlinear, $A(s,t)$ can develop singularities which are called *focal points* of S. Let us see some examples.

Example 2.3 Let $S_t = \partial B_t(p)$, where $B_t(p) = \{x \in M : \ |x,p| < t\}$ is the Riemannian ball.

Then V is radial and

$$A(t) \sim \frac{1}{t} I \quad \text{as } t \to 0 \tag{2.9}$$

because a Riemannian manifold behaves as a Euclidian space for $t \to 0$.

Example 2.4 ([9]) Let us suppose that $R_V = kI$, $k \in \mathbb{R}$, that is M has constant curvature. Moreover let us suppose that $A = aI$, where a is a real function defined on M (A is the second fundamental form of a family of *umbilical* hypersurfaces). In this case equation (2.7) becomes:

$$a' + a^2 + k = 0.$$

If $k > 0$, then M is a sphere (if it is assumed to be complete and simply connected). The solutions are given by

$$a(t) = \sqrt{k} \cot(\sqrt{k}(t - t_0)).$$

This corresponds to the fact that there is (up to congruence) only one equidistant family of umbilical hypersurfaces in the sphere, namlely concentric Riemannian spheres (latitude circles).

Fig. 3.

If $k = 0$, then M is a Euclidean space and the solutions are

$$a(t) = \frac{1}{(t - t_0)}, \qquad a(t) = 0.$$

Fig. 4.

These solutions correspond to the three umbilical parallel hypersurface families in euclidean space: concentric spheres with increasing ($t > t_0$) or decreasing ($t < t_0$) radii and parallel hyperplanes.
Finally, if $k < 0$, the space M is hyperbolic. The solutions are given by

$$a(t) = \sqrt{|k|}\coth(\sqrt{|k|}(t - t_0), \qquad a(t) = \sqrt{|k|}\tanh(\sqrt{|k|}(t - t_0), \qquad a(t) = \pm\sqrt{|k|}.$$

These solutions correspond to the five families of equidistant hypersurfaces in the hyperbolic space: Concentric spheres with increasing ($t > t_0$) or decreasing ($t < t_0$) radii, hypersurfaces which are parallel to an $(n-1)$-dimensional hyperbolic subspace, and expanding ($t > t_0$) or contracting ($t < t_0$) horospheres.

Fig. 5.

3. Comparison theory.

We want to derive a comparison theorem for solutions of the Riccati equation $A' + A^2 + R_V = 0$ (cf. 2.7). Fixing an integral curve γ of V (which is a geodesic) and identifying all tangent spaces $T_{\gamma(t)}M$ by parallel displacement (i.e. via an orthonormal basis $(E_i(t))$ of vector fields along γ which are *parallel*, i.e. $E_i' = 0$), we consider $A(t)$ as a self adjoint endomorphism on a single vector space $E = T_{\gamma(0)}M$. More generally, let E be a finite-dimensional real vector space with euclidean inner product $\langle\ ,\ \rangle$. The space $S(E)$ of self adjoint endomorphisms inherits the inner product

$$\langle A, B\rangle = \mathrm{trace}(A \cdot B) \tag{3.1}$$

for $A, B \in S(E)$. We get a partial ordering $\leq$ on $S(E)$ by putting $A \leq B$ if $\langle Ax, x\rangle \leq \langle Bx, x\rangle$ for every $x \in E$.

Theorem 3.1 (cf. [14], [9]) Let $R_1, R_2 : \mathbb{R} \to S(E)$ be smooth with $R_1 \geq R_2$. For $i \in \{1, 2\}$ let $A_i : [t_0, t_i) \to S(E)$ be a solution of

$$A_i' + A_i^2 + R_i = 0 \tag{3.2}$$

with maximal $t_i \in (t_0, \infty]$. Assume that $A_1(t_0) \leq A_2(t_0)$. Then $t_1 \leq t_2$ and $A_1(t) \leq A_2(t)$ on (t_0, t_1).

Proof. Let $U = A_2 - A_1$; then $U(t_0) \geq 0$ and

$$U' = A_2' - A_1' = A_1^2 - A_2^2 + R_1 - R_2. \tag{3.3}$$

We define $S = R_1 - R_2 \geq 0$ and $X = -\frac{1}{2}(A_1 + A_2)$; the equation (3.3) takes the form

$$U' = XU + UX + S. \tag{3.4}$$

We solve (3.4) by the variation of constant method (see [14], pag. 211, Remark 1). Let $t' = \min\{t_1, t_2\}$ and $g : (t_0, t') \to S(E)$ be a non-singular solution of the homogeneous equation

$$g' = Xg. \tag{3.5}$$

Now a solution U of (3.4) is obtained as

$$U = g \cdot V \cdot g^T$$

where V verifies

$$V' = g^{-1} \cdot S \cdot (g^{-1})^T. \tag{3.6}$$

¿From $S \geq 0$ we get $V' \geq 0$; this, combined with $V(0) \geq 0$, implies that $V \geq 0$ and hence $U \geq 0$. Thus $A_1 \leq A_2$ on (t_0, t'). Since A_i' is bounded from above, a singularity can only be negative (going to $-\infty$). So $A_1 \leq A_2$ implies $t' = t_1 \leq t_2$. ∎

Remark 3.2 Theorem 3.1 still holds if A_1, A_2 are singular at t_0, but $U = A_2 - A_1$ has a continuous extension to 0 with $U(0) \geq 0$. See [14] for the proof. A similar argument also shows that $t_1 < t_2$ if $A_1(t_0) < A_2(t_0)$; for a different proof of this fact see [11], Lemma 3.1.

The geometric interpretation of Theorem 3.1 is: principal curvatures (i.e. eigenvalues of the shape operator) of equidistant hypersurfaces decrease faster on the space of larger curvature. In particular, this is true for Riemannian spheres, as follows by Remark 3.2).

Now we want to find a comparison theorem for equation (2.6). For $A \in S(E)$, denote by $\lambda_-(A)$ the lowest eigenvalue and by $\lambda_+(A)$ the highest eigenvalue of A.

Theorem 3.3 Let $A_1, A_2 : (t_0, t') \to S(E)$ such that

$$\lambda_+(A_1(t)) \leq \lambda_-(A_2(t)) \quad \text{everywhere.} \tag{3.7}$$

Let $J_1, J_2 : (t_0, t') \to E$ be nonzero solutions of $J_i' = A_i \cdot J_i$. Then $\|J_1\|/\|J_2\|$ is monotoneously decreasing.
Moreover, if

$$\lim_{t \searrow t_0} \frac{\|J_1\|}{\|J_2\|}(t) = 1, \tag{3.8}$$

then $\|J_1\| \leq \|J_2\|$.
Equality holds at some $t \in (t_0, t')$ iff for $i = 1, 2$ we have $J_i = j \cdot v_i$ on $[t_0, t]$ for some constant vector $v_i \in E$ with $Av_i = \lambda \cdot v_i$ and $j' = \lambda \cdot j$, where $\lambda = \lambda_+(A_1) = \lambda_-(A_2)$.

Proof. Since $\|J_i\|$ is smooth, we can consider

$$\frac{\|J_i\|'}{\|J_i\|} = \frac{< J_i', J_i >}{\|J_i\|^2} = \frac{< A_i J_i, J_i >}{< J_i, J_i >} \in [\lambda_-(A_i), \lambda_+(A_i)]$$

so that

$$\log(\|J_1\|)' = \frac{\|J_1\|'}{\|J_1\|} \leq \lambda_+(A_1) \leq \lambda_-(A_2) \leq \frac{\|J_2\|'}{\|J_2\|} = \log(\|J_2\|)',$$

hence

$$\left(\log \frac{\|J_1\|}{\|J_2\|}\right)' \leq 0$$

which implies that $\|J_1\|/\|J_2\|$ is monotoneously decreasing.
If $\|J_1\|/\|J_2\|$ has the same value 1 at t_0 and t, then $\|J_1\| = \|J_2\|$ on $[t_0, t]$ and we recieve $J_i' = A_i J_i = \lambda J_i$ from which the conclusion follows.

$\blacksquare$

We consider the most important special cases due to Rauch and Berger (called Rauch I and Rauch II in [5]):

Rauch I
Suppose that J_i for $i = 1, 2$ are solutions of $J_i'' + R_i J_i = 0$ with $\lambda_-(R_1) \geq \lambda_+(R_2)$ and

$$J_i(0) = 0, \ \|J_1'(0)\| = \|J_2'(0)\|.$$

Then $\|J_1\| \leq \|J_2\|$ up to the first zero of J_1.

Rauch II
Suppose that J_i for $i = 1, 2$ are solutions of $J_i'' + R_i J_i = 0$ with $\lambda_-(R_1) \geq \lambda_+(R_2)$ and

$$J_i'(0) = 0, \ \|J_1(0)\| = \|J_2(0)\|.$$

Then $\|J_1\| \leq \|J_2\|$ up to the first zero of J_1.

In fact we apply the theorems 3.1 and 3.3 where in the first case, $A_i(t) \sim t^{-1} I$ as $t \to 0$ and in the second case, $A_i(0) = 0$.

Corollary 3.4 Let M be a complete manifold with $K \geq 0$, $p_0, p_1 \in M$ and $\gamma : [0, 1] \to M$ a shortest geodesic segment connecting p_0 and p_1. Let $X \perp \gamma'$ be a parallel vector field along γ. Put $p_s(t) = \exp tX(s)$ for all $s \in [0, 1]$. Then

$$|p_0(t), p_1(t)| \leq |p_0, p_1|$$

with equality for some $t > 0$ only if $p_0, p_1, p_1(t), p_0(t)$ bound a flat totally geodesic rectangle.

Proof. We have

$$|p_0(t), p_1(t)| \leq \int_0^1 \|\frac{\partial}{\partial s} p_s(t)\| dt$$

and $J_s(t) = \frac{\partial}{\partial s} p_s(t)$ is a Jacobi field along the geodesic $\gamma_s(t) = p_s(t)$ with $J_s'(0) = 0$. Thus comparing with the euclidean case we get from Rauch II that $\|J_s(t)\| \leq \|J_s(0)\|$ which shows the inequality. If we have equality at $t_1 > 0$, the equality discussion of Theorem 3.3 shows that J_s is parallel along $\gamma_s|[0, t_1]$. Moreover, the curves $s \mapsto p_s(t)$ are shortest geodesics of constant length for $0 \leq t \leq t_1$. Thus the surface $p : (s, t) \mapsto p_s(t)$ is a flat rectangle in M with

$$\frac{D}{\partial s}\frac{\partial p}{\partial s} = \frac{D}{\partial t}\frac{\partial p}{\partial s} = \frac{D}{\partial t}\frac{\partial p}{\partial t} = 0,$$

so it is also *totally geodesic*, i.e. covariant derivatives of vector fields tangent to p remain tangent to p. ■

4. Average comparison theorems.

Now we consider the trace of the Riccati equation $A' + A^2 + R_V = 0$ for self adjoint A. Since trace and derivative commute, we get

$$\operatorname{trace}(A)' + \operatorname{trace}(A^2) + \operatorname{Ric}(V) = 0. \tag{4.1}$$

This is unfortunately not a differential equation for trace(A), because of the term trace(A^2). However, put

$$a = \frac{\operatorname{trace}(A)}{n-1}.$$

(Note that $A(V) = D_V V = 0$, so we consider A as an endomorphism on the $(n-1)$-dimensional subspace $E = V^\perp$ of the tangent space.) Then

$$A = aI + A_0,$$

with $\operatorname{trace}(A_0) = 0$, so A_0 and I are perpendicular. Hence,

$$\operatorname{trace}(A^2) = \|A\|^2 = a^2\|I\|^2 + \|A_0\|^2 = (n-1)a^2 + \|A_0\|^2$$

and we get, from the trace equation (4.1):

$$a' + a^2 + r = 0 \tag{4.2}$$

with

$$r = \frac{1}{n-1}\left(\|A_0\|^2 + Ric(V)\right) \geq \frac{1}{n-1} Ric(V).$$

Geometric meaning: $a(t)$ is the mean curvature of S_t.

Theorem 4.1 Suppose that $A : [t_0, t_1) \to S(E)$ ($t_1 \leq +\infty$ maximal) is a solution of

$$A' + A^2 + R = 0 \tag{4.3}$$

where $R : \mathbb{R} \to S(E)$ is given; suppose that for some constant $k \in \mathbb{R}$:

(1) $\operatorname{trace}(R) \geq (n-1)k$
(2) $\operatorname{trace}(A(t_0)) \leq (n-1)\overline{a}(t_0)$
where $\overline{a} : [t_0, t_2) \to \mathbb{R}$ is a solution of

$$\overline{a}' + \overline{a}^2 + k = 0 \tag{4.4}$$

with $t_2 \leq +\infty$ maximal. Let

$$a = \frac{\text{trace}(A)}{n-1}. \tag{4.5}$$

Then $t_1 \leq t_2$ and $a(t) \leq \overline{a}(t)$ for $t \in [t_0, t_1)$.

Proof. Apply theorem 3.1 with (R_1, A_1, R_2, A_2) replaced with $(r, a, k, \overline{a})$.

Remark 4.2 By Remark 3.2, the theorem remains true if $A(t) \sim \frac{1}{t-t_0} I$ and $\overline{a}$ is the solution of (4.4) with a pole at t_0, i.e. $\overline{a} = s'/s$, where s is the solution of

$$s'' + ks = 0, \quad s(t_0) = 0, \quad s'(t_0) = 1.$$

Next, let $J_1, \ldots, J_{n-1}$ be a basis of solutions of $J' = A \cdot J$, and put

$$j = det(J_1, \ldots, J_{n-1}).$$

Since

$$(J_1 \wedge \ldots \wedge J_{n-1})' = \sum_{k=1}^{n-1} J_1 \wedge \ldots \wedge A \cdot J_k \wedge \ldots \wedge J_{n-1},$$

we get

$$j' = (n-1)a \cdot j. \tag{4.6}$$

Geometrically, equation (4.6) says how the volume element of S_t, namely $\det(d\phi_t)$ (see page 9 of chapter 2), changes with t.

Theorem 4.3 Let $A : [t_0, t_1) \to S(V)$ be given with

$$a \leq \overline{a},$$

where $a = \frac{1}{n-1}\text{trace}(A)$, and let j be as above. Choose $\overline{j}$ such that

$$\overline{j}' = (n-1)\overline{a} \cdot \overline{j}.$$

Then $j/\overline{j}$ is monotonously decreasing.

Proof. Apply theorem 3.3 with (A_1, J_1, A_2, J_2) replaced with $((n-1)a, j, (n-1)\overline{a}, \overline{j})$.

5. Bishop - Gromov inequality

Let M be a complete connected Riemannian manifold. By the theorem of Hopf and Rinow (cf. [29]), any two points $p, q \in M$ can be connected by a *shortest* geodesic γ, i.e. $L(\gamma) = |p, q|$. Let $S_pM = \{v \in T_pM : \ \|v\| = 1\}$ be the unit sphere in T_pM. For any $v \in S_pM$, we define

$$cut(v) = \max\{t \ : \ \gamma_v|_{[0,t]} \text{ is shortest}\}.$$

This defines a function $cut : S_pM \to (0, \infty]$, the *cut locus distance*, which is continuous (cf [5], p.94). Let

$$C_p = \{tv \ : \ v \in S_pM, \quad t \leq cut(v)\}. \tag{5.1}$$

This is a closed subset of T_pM, and its boundary ∂C_p (sometimes also $\exp_p(\partial C_p) \subset M$) is called the *cut locus* of the point p. It follows from this definition that

$$B_r(p) = \exp_p(B_r(0)) = \exp_p(B_r(0) \cap C_p) \quad \forall r > 0. \tag{5.2}$$

In fact, if we choose $q \in B_r(p)$, there exists a shortest geodesic γ_v joining p and q; the length of γ_v should be $\leq cut(v)$, hence $v \in C_p$ (theorem of Hopf - Rinow).

Example 5.1 On the unit sphere we have $cut(v) = \pi$ for every v. In fact, in every direction, the geodesic is a meridian, hence it is shortest up to the opposite ("antipodal") point.

Example 5.2 On the cylinder $S^1 \times \mathbb{R}$, we have $cut(v) = \pi / \cos\alpha$ where α is the angle between v and the S^1-direction.

Fig. 6.

There are two ways how a geodesic $\gamma = \gamma_v : [0, \infty) \to M$ (where $v \in S_pM$) can cease to be shortest beyond the parameter $t_0 = cut(v)$ (cf. [5], p.93): Either there exists a nonzero Jacobi field J along γ which vanishes at 0 and t_0 - in this case, $\gamma(t_0)$ is called a *conjugate point* of p (cf. Example 5.1), or there exists a second geodesic $\sigma \neq \gamma$ of the same length which also connects p and $\gamma(t_0)$ (cf. Example 5.2). Hence $q = \gamma(t_0)$ is in the cut locus of $p = \gamma(0)$ iff p is in the cut locus of q. Moreover, there are no conjugate points on $\gamma|[0, cut(v))$. The conjugate points in turn are the singular values of the exponential map $\exp_p$; more precisely, we have:

Lemma 5.3 Let $J(t)$ be the Jacobi field along γ_v defined by $J(0) = 0$, $J'(0) = w$. Then we have

$$d(\exp_p)_{tv}.tw = J(t).$$

In particular, $d(\exp_p)_{tv}$ is singular if and only if $\exp_p(tv)$ is a conjugate point of p.

Proof. Let $w \in T_vT_pM \equiv T_pM$. Then we have

$$d(\exp_p)_v.w = \left.\frac{d}{ds}\right|_{s=0} \exp_p(v + sw) = \left.\frac{d}{ds}\right|_{s=0} \gamma_{v+sw}(1). \tag{5.3}$$

If we let

$$J(t) = \left.\frac{\partial}{\partial s}\right|_{s=0} \gamma_{v+sw}(t), \tag{5.4}$$

then J is the Jacobi field along γ_v with initial conditions $J(0) = 0$ and

$$\begin{aligned} J'(0) &= \left.\frac{D}{\partial t}\right|_0 \left.\frac{\partial}{\partial s}\right|_0 \gamma_{v+sw}(t) \\ &= \left.\frac{D}{\partial s}\right|_0 \left.\frac{\partial}{\partial t}\right|_0 \gamma_{v+sw}(t) . \\ &= \left.\frac{D}{\partial s}\right|_0 (v + sw) = w \end{aligned}$$

Therefore we get

$$d(\exp_p)_v \cdot w = J(1), \tag{5.5}$$

and generally

$$d(\exp_p)_{tv} \cdot tw = J(t). \tag{5.6}$$

■

Remark 5.4 Consequently, on the interior of C_p, the exponential map $\exp_p$ is injective and regular, hence a diffeomorphism. Note that $\mathrm{Int}(C_p)$ is star-shape, thus it is contractive; hence also its image is contractive. But by Hopf-Rinow, the whole manifold M is the image of $\exp_p : C_p \to M$, so the topology of M is given by the image of the boundary ∂C_p.

After these preparations, we come to the main theorem of this section.

Theorem 5.5 Let us consider a manifold M^n with Ricci curvature satisfying

$$\frac{Ric}{n-1} \geq k.$$

Let $\bar{M}$ be the complete simply connected n-manifold with constant curvature k (*standard space of constant curvature* k) and $\bar{B}_r \subset \bar{M}$ the ball of radius r in $\bar{M}$. Then, for all $p \in M$, we have that

$$\frac{\mathrm{Vol}B_r(p)}{\mathrm{Vol}\bar{B}_r} \searrow_r \tag{5.7}$$

i.e. this quotient is monotonely decresing with r. Moreover, for $r \to 0$, the quotient goes to one.

Corollary 5.6 For any two positive real numbers $R > r$ we have

$$\frac{\mathrm{Vol}B_R(p)}{\mathrm{Vol}B_r(p)} \leq \frac{\mathrm{Vol}\bar{B}_R}{\mathrm{Vol}\bar{B}_r}. \tag{5.8}$$

Remark 5.7 Corollary 5.6 gives an upper bound for the growth of the metric balls in M. Moreover, if equality holds for some $r < R$, then $B_R(p)$ is isometric to $\bar{B}_r$ (this can be seen from the proof).

Proof of the theorem. By (5.3) we have

$$\mathrm{Vol}B_r(p) = \int_{B_r(0)\cap C_p} \det\left(d(\exp_p)_u\right) du. \tag{5.9}$$

Passing to polar coordinates and denoting $r(v) = \min\{r, cut(v)\}$, we get

$$\mathrm{Vol}B_r(p) = \int_S \int_0^{r(v)} \det\left(d(\exp_p)_{tv}\right) t^{n-1} dt\, dv \tag{5.10}$$

where $S := S_1(0) \subset T_pM$. If we consider a basis $e_1, \ldots, e_{n-1}$ of $v^\perp \subset T_pM$, then by Lemma 5.3,

$$d(\exp_p)_{tv} e_i = \frac{1}{t} d(\exp_p)_{tv} t e_i = \frac{1}{t} J_i(t),$$

where J_i is the Jacobi field along γ_v with $J_i(0) = 0$ and $J_i'(0) = e_i$. Hence

$$\det\left(d(\exp_p)_{tv}\right) = \frac{1}{t^{n-1}} \det\left(J_1(t), \ldots, J_{n-1}(t)\right), \tag{5.11}$$

and equation (5.10) becomes

$$\mathrm{Vol} B_r(p) = \int_S \int_0^{r(v)} j_v(t) dt\, dv \tag{5.12}$$

where

$$j_v(t) = \det\left(J_1(t), \ldots, J_{n-1}(t)\right). \tag{5.13}$$

If we put $j_v(t) = 0$ for $t > cut(v)$, then by the comparison theorem 4.3 we get

$$(j_v/\bar{j}) \searrow$$

on $[0, r]$ and hence

$$q := \frac{1}{\mathrm{Vol}(S)} \int_S (j_v/\bar{j}) dv$$

is still monotone. Moreover,

$$\mathrm{Vol}\bar{B}_r = \int_S \int_0^r \bar{j}(t) dt dv = \mathrm{Vol}(S) \int_0^r \bar{j}(t) dt. \tag{5.14}$$

Therefore we have that

$$\frac{\mathrm{Vol} B_r(p)}{\mathrm{Vol}\bar{B}_r} = \frac{\int_0^r q(t)\bar{j}(t) dt}{\int_0^r \bar{j}(t) dt} \tag{5.15}$$

is a monotone decreasing function in r, because the mean of a monotone function on growing intervals is still monotone.

If $r \to 0$, both volumes approximate the euclidean ball volume, hence the quotient goes to one.

6. Toponogov's Triangle Comparison Theorem

Let us fix $o \in M$ and let $\rho = |o, \cdot|$. We already know that near o, precisely in $\exp_o(Int(C_o)\backslash\{0\})$, ρ is a C^∞ function and

$$\rho(\exp_o(v)) = \|v\|. \tag{6.1}$$

Let us consider the unit radial field $V = \nabla\rho$. Then $S_r = \partial B_r(o)$ is a family of equidistant hypersurfaces, as in chapter 2.

Suppose that the sectional curvature K of M is $\geq k$. If $\tilde{M}$ is the standard space of sectional curvature k, then, by the comparison theorem 3.1, we get

$$A \leq \tilde{A} = \frac{s'}{s} I, \tag{6.2}$$

where s is a solution of $s'' + ks = 0$ with initial dates $s(0) = 0, s'(0) = 1$, and $A = DV = D\nabla\rho$ is the Hessian of ρ. (Recall from Example 2.4 that $a = s'/s$ is the (unique) solution of the equation $a' + a^2 + k = 0$ with a pole at $t = 0$.)

Therefore,

$$D\nabla\rho|_{V^\perp} \leq \frac{s'}{s} I, \tag{6.3}$$

while

$$D\nabla\rho|_{\mathbb{R}V} = 0, \tag{6.4}$$

because ρ grows linearly along the integral curves of V. Analogous relations hold for $\tilde{\rho}$:

$$D\nabla\tilde{\rho}|_{V^\perp} = \frac{s'}{s} I \tag{6.5}$$

$$D\nabla\tilde{\rho}|_{\mathbb{R}V} = 0 \tag{6.6}$$

Now we want to find a unique estimate for the whole Hessian. To get this we modify ρ (and analogously $\tilde{\rho}$) suitably: Consider $\sigma = f \circ \rho$, where $f : \mathbb{R} \to \mathbb{R}$ is a function yet to be determined. Then

$$D\nabla(f \circ \rho) = D\left((f' \circ \rho)\nabla\rho\right) = f''(\rho) d\rho \cdot \nabla\rho + f'(\rho) D\nabla\rho. \tag{6.7}$$

On $V^\perp$ we have $f''(\rho)d\rho \cdot \nabla\rho = 0$ and $D\nabla\rho \leq (s'/s)I$; while $f''(\rho)d\rho \cdot \nabla\rho = f''(\rho)I$ and $D\nabla\rho = 0$ on $\mathbb{R} \cdot V$. If we choose f as a principal function of s, i.e. $f' = s$, then we get

$$f''(\rho) = -kf(\rho) + C,$$

hence (6.3) and (6.4) give

$$D\nabla\sigma \leq -k\sigma I + C \tag{6.8}$$

where C is some fixed constant. Analogously, for $\tilde{\sigma} = f \circ \tilde{\rho}$, we get from (6.5) and (6.6)

$$D\nabla\tilde{\sigma} = -k\tilde{\sigma}I + C. \tag{6.9}$$

Theorem 6.1 (*Toponogov's triangle comparison theorem*) [18], [5], [24]
Let M be a complete Riemannian manifold with sectional curvature $K \geq k$. Let $\tilde{M}$ be the standard space of constant curvature k. Let $p_0,\ p_1,\ o \in M$, and choose corresponding points $\tilde{p}_0,\ \tilde{p}_1,\ \tilde{o} \in \tilde{M}$. Let γ be a geodesic from p_0 to p_1, and β_i a shortest geodesic from p_i to o, $i = 0, 1$, all parametrized by arc length, and let $\tilde{\gamma}, \tilde{\beta}_i$ be the corresponding curves in $\tilde{M}$, with $L(\gamma) = L(\tilde{\gamma}) = L$ and $L(\beta_i) = L(\tilde{\beta}_i)$. Let us suppose that all the lengths are smaller than $\pi/\sqrt{k}$, if $k > 0$. Then we have

$$|o, \gamma(t)| \leq |\tilde{o}, \tilde{\gamma}(t)| \quad \forall t \in [0, L]. \tag{6.10}$$

Fig. 7.

Remark 6.2 The hypothesis made on lengths (when $k > 0$) implies that the geodesics in $\tilde{M}$ are shortest.

Corollary 6.3 Let $\alpha_0 = \angle(\beta_0', \gamma'(0))$, $\alpha_1 = \angle(\beta_1', -\gamma'(L))$ and let $\tilde{\alpha}_0,\ \tilde{\alpha}_1$ the corresponding angles in $\tilde{M}$. Then

$$\alpha_i \geq \tilde{\alpha}_i. \tag{6.11}$$

Proof of the corollary. Let us suppose, by contradiction, that $\alpha_0 < \tilde{\alpha}_0$. Suppose first that p_0 is not in the cut locus of o. Then $\exp_o$ is invertible near p_0 (cf. Lemma 5.3).

Let β_t be the shortest geodesic joining o to $\gamma(t)$; the corresponding one $\tilde{\beta}_t$ is a shortest geodesic in $\tilde{M}$ (for t close to 0), hence

$$L(\beta_t) \geq |o, \gamma(t)|,$$

$$L(\tilde{\beta}_t) = |\tilde{o}, \tilde{\gamma}(t)|.$$

We have

$$L(\tilde{\beta}_t) = |o, p_0| + t \left. \frac{d}{dt} \right|_0 L(\tilde{\beta}_t) + O(t^2) \tag{6.12}$$

$$L(\beta_t) = |o, p_0| + t \left. \frac{d}{dt} \right|_0 L(\beta_t) + O(t^2) \tag{6.13}$$

and, by the first variation formula for curves (cf. [5], p.5), we get

$$\left. \frac{d}{dt} \right|_0 L(\beta_t) = - < \gamma'(0), \beta_0'(0) >$$

$$\left. \frac{d}{dt} \right|_0 L(\tilde{\beta}_t) = - < \tilde{\gamma}'(0), \tilde{\beta}_0'(0) > .$$

Since we supposed $\alpha_0 < \tilde{\alpha}_0$, for small t we get $L(\gamma_t) < L(\tilde{\gamma}_t)$, which implies

$$|o, \gamma(t)| \leq L(\gamma_t) < L(\tilde{\gamma}_t) = |\tilde{o}, \tilde{\gamma}(t)|.$$

Thus, by Toponogov's theorem, we get a contradiction.

If p_0 happens to be a cut locus point of o, we choose $o_\varepsilon = \beta_0(\varepsilon)$ on β_0 close to o. Then certainly p_0 is not in the cut locus of o_ε. Now we put β_t the broken geodesic $\beta|[0, \varepsilon] \cup \beta_{\varepsilon,t}$ where $\beta_{\varepsilon,t}$ denotes the shortest geodesic from o_ε to $\gamma(t)$, and the same argument holds.

Proof of theorem 6.1

Let us define $\rho = |o, \cdot|$, $\tilde{\rho} = |\tilde{o}, \cdot|$, and $\sigma = f \circ \rho$, $\tilde{\sigma} = f \circ \tilde{\rho}$. Consider the function

$$\delta = \sigma \circ \gamma - \tilde{\sigma} \circ \tilde{\gamma}. \tag{6.14}$$

Hence we have to prove that

$$\delta \geq 0 \qquad \text{on } [0, L]. \tag{6.15}$$

Fig. 8.

We prove (6.15) by contradiction. Suppose that there is $t \in [0, L]$ such that $\delta(t) < 0$, and let $m = \min_{[0,L]} \delta(t) < 0$.
We choose $k' > k$ sufficiently close to k and $\tau > 0$ such that

$$L < \frac{\pi}{\sqrt{k'}} - \tau. \tag{6.16}$$

It is easy to find a solution a_0 of the equation $a_0'' + k'a_0 = 0$, with $a_0(-\tau) = 0$ and $a_0|_{[0,L]} \leq m$. Then there exists $\lambda > 0$ such that $a = \lambda a_0$ satisfies the following properties:

1. $a \leq \delta$
2. $a(t_0) = \delta(t_0)$ for some $t_0 \in (0, L)$.

Case 1: $\gamma(t_0)$ is not a cut locus point of o. Thus δ is of class C^∞ in a neighborhood of t_0 and

$$(\sigma \circ \gamma)'' =< D_{\gamma'} \nabla \sigma, \gamma' >\leq -k(\sigma \circ \gamma) + C, \tag{6.17}$$

where the inequality follows from (6.8). By eqution (6.9) we get

$$(\tilde{\sigma} \circ \tilde{\gamma})'' =< D_{\tilde{\gamma}'} \nabla \tilde{\sigma}, \tilde{\gamma}' >= -k(\tilde{\sigma} \circ \tilde{\gamma}) + C. \tag{6.18}$$

Hence

$$\delta'' \leq -k\delta. \tag{6.19}$$

On the other hand $a'' = -k'a$. Moreover, in t_0 we have $\delta(t_0) = a(t_0) < 0$, which implies

$$(\delta - a)''(t_0) \leq \delta(t_0)(k' - k) < 0. \tag{6.20}$$

This is a contradiction because $\delta - a$ takes a minimum at t_0.

Case 2: $\gamma(t_0)$ is a cut locus point of o. Let β be a shortest geodesic from o to $\gamma(t_0)$. We choose o_ε on β close to o, say $|o_\varepsilon, o| = \varepsilon$. Then we replace ρ by $\rho_\varepsilon(x) := |x, o_\varepsilon| + |o_\varepsilon, o|$. By triangle inequality,

$$\rho_\varepsilon(x) \geq \rho(x), \tag{6.21}$$

and equality holds at $x = \gamma(t_0)$. In other words, ρ_ε is an *upper support function of* ρ at $\gamma(t_0)$. Since β is shortest from o to $\gamma(t_0)$, o_ε is not a cut point of $\gamma(t_0)$, and therefore, $\gamma(t_0)$ is not a cut point of o_ε (cf. Ch.5). Putting $\sigma_\varepsilon = f \circ \rho_\varepsilon$, we get the same estimates as in Case 1 for σ_ε in place of σ, up to a small error which goes to zero as $\varepsilon \to 0$:

$$(\sigma_\varepsilon \circ \gamma)'' \leq -k(\sigma_\varepsilon \circ \gamma) + C + error. \tag{6.22}$$

Now σ_ε is an upper support function of σ at $\gamma(t_0)$ as f is monotoneously increasing. Hence $\delta_\varepsilon - a$ is an upper support function of $\delta - a$ at t_0 where $\delta_\varepsilon = \sigma_\varepsilon - \tilde{\sigma}$. Thus it also takes a minimum at t_0. But this is a contradiction, because $(\delta_\varepsilon - a)''(t_0) < 0$ by (6.20).
■

Remark 6.4 The above proof is essentially due to Karcher ([24]). Recently, M. Kürzel ([27]) extended this proof to the case where curvature bounds are given which depend radially on the point (rather than being constant).

7. Number of generators and growth of the fundamental group

Let M be a complete Riemannian manifold and $\hat{M}$ its universal covering. The *fundamental group* $\pi_1(M)$ will be viewed as group of deck transformations acting on $\hat{M}$. In other words, M is the orbit space of a discrete group $\Gamma \cong \pi_1(M)$ of isometries of $\hat{M}$ *acting freely* on $\hat{M}$, i.e. if $g \in \Gamma$ with $g(p) = p$ for some $p \in M$, then $g = 1$.

Remark 7.1 The fundamental group of any compact Riemannian manifold M is finitely generated.

Proof. There exists a compact fundamental domain F (see definition below) for the action of Γ on $\hat{M}$; e.g. one may take the so called Dirichlet fundamental domain

$$F = \{x \in \hat{M} \ ; \ \ |x, o| \leq |x, go| \ \forall g \in \Gamma\}.$$

We say that $g \in \Gamma$ is *small* if $gF \cap F \neq \emptyset$, i.e. if the fundamental domains F and gF are neighbors. If $d(F)$ denotes the *diameter* of F, i.e. the largest possible distance within F, then $gF \subset B_{2d(F)}(o)$ for all small g, for some fixed $o \in F$. Since the subsets $g(Int(F))$ are all disjoint with equal volume, there can be only finitely many of them in this ball, hence there exist only finitely many small g. We claim that they form a set of generators. In fact, let $g \in \Gamma$ arbitrary. Choose a geodesic segment γ from o to go. Then γ is covered by finitely many fundamental domains $g_0F, ..., g_NF$ where $g_0 = 1$ and $g_N = g$, and $g_{i-1}F$, g_iF are neighbors. Thus $g_i^{-1}g_{i-1}$ is small, and hence g is a composition of small group elements. ■

Definition 7.2 A closed subset $F \subset \hat{M}$ is a *fundamental domain* for a group Γ acting on $\hat{M}$ if
(a.) $Int(F) \cap Int(gF) = \emptyset \ \forall g \neq 1$;
(b.) $\Gamma \cdot F = \hat{M}$.

For a noncompact manifold M, the fundamental group may have infinitely many generators. The next theorem shows that this does not happen if M has $K \geq 0$; in fact, there is an a-priori bound on the *number of generators*, i.e. the cardinality of a suitably chosen set of generators:

Theorem 7.3 (*Gromov 1978*, cf [24])
There exists a number $c(n)$ such that:
(a) the number of generators for $\pi_1(M)$ is $\leq c(n)$ for any n-dimensional complete manifold M with curvature $K \geq 0$.
(b) the number of generators for $\pi_1(M)$ is $\leq c(n)^{1+kD}$ for any n-dimensional compact manifold M with curvature $K \geq -k^2$ and diameter bounded, $\mathrm{diam}(M) \leq D$.

Proof. We prove only part a); the second part is similar, but more technical (see Remark at the end of the proof).
We define a "norm" in Γ as follows:

$$|g| = |p, g(p)|$$

for some fixed $p \in \hat{M}$. There exists $g_1 \in \Gamma \setminus \{1\}$ with $|g_1|$ minimal (not necessarily unique). By induction, we can construct a sequence (g_j): given $g_1, \ldots, g_k$, we define

$$\Gamma_k = \langle g_1, \ldots, g_k \rangle \subset \Gamma$$

and choose $g_{k+1} \in \Gamma \setminus \Gamma_k$ such that $|g_{k+1}|$ has minimum norm in $\Gamma \setminus \Gamma_k$. To finish the proof, we only have to show

Claim: $\Gamma_k = \Gamma$ for some $k \leq c_0(n) := 2\sqrt{5}^n$.

Proof of the claim: for $j > i$ we have $|g_j| \geq |g_i|$, and moreover

$$|g_i(p), g_j(p)| = |p, g_i^{-1} g_j(p)| = |g_i^{-1} g_j| \geq |g_j|$$

since $g_i^{-1} g_j \in \Gamma \setminus \Gamma_{j-1}$ (otherwise $g_i^{-1} g_j,\ g_i \in \Gamma_{j-1}$ which would imply that $g_j \in \Gamma_{j-1}$ contradicting the choice of g_j). Now consider the triangle p, $p_i = g_i(p)$, $p_j = g_j(p)$.

Fig. 9.

Let γ_{v_i} be a shortest geodesic from p to p_i, and α_{ij} the angle between v_i and v_j. The standard space $\tilde{M}$ of zero curvature is euclidean space $\mathbb{R}^n$. Considering the comparison triangle $\tilde{p}$, $\tilde{p}_i$, $\tilde{p}_j$ in $\tilde{M}$, we have $\tilde{\alpha}_{ij} \geq 60°$ for the corresponding angle $\tilde{\alpha}_{ij}$ in $\tilde{p}$. (Note that $\tilde{\alpha}_{ij}$ is opposite to the largest edge in that triangle.) By Toponogov's theorem then

$$\alpha_{ij} \geq 60° \quad \forall i \neq j. \tag{7.1}$$

Fig. 10.

There are at most $2\sqrt{5}^n$ vectors that satisfy (7.1). Namely, for any two vectors v_i, v_j of this kind, balls of radii $\frac{1}{2}$ are disjoint and their inner half balls are contained in $B_{\sqrt{5}/2}(0)$, as the figure shows.

Thus, if there are k such vectors, then

$$\text{Vol } B_{\sqrt{5}/2}(0) \geq \frac{k}{2} \text{Vol } B_{\frac{1}{2}}$$

hence

$$k \leq 2 \frac{\text{Vol } B_{\frac{\sqrt{5}}{2}}}{\text{Vol } B_{\frac{1}{2}}} = 2\sqrt{5}^n.$$

This finishes the proof of the claim and of the theorem.

Remark 7.4 A much better (but more difficult) estimate was given by U.Abresch (cf. [1]).

Remark 7.5 In Case (b), we use comparison with a hyperbolic triangle (curvature $-k^2$) instead of a euclidean one. Since the side lengths are a priori bounded by the diameter bound D, this is not much difference. To see that such a bound is necessary, let M be a compact surface of genus g with constant negative curvature. Then $\pi_1(M_g)$ is generated by $2g$ elements, hence is not bounded as $g \to \infty$). Nevertheless the theorem holds, since either the curvature or the diameter are unbounded as $g \to \infty$.

Now let us assume that M has $Ric \geq 0$ rather than $K \geq 0$. If M is complete and noncompact, it is an open question whether the fundamental group is finitely generated. However, for any finitely generated subgroup, the growth of this group is only polynomial:

Definition 7.6 Let Γ be a finitely generated group and G a finite set of generators of Γ with $G = G^{-1}$ and $1 \in G$. We define the *growth function* $N(k)$ (depending on Γ and G) as follows:

$$N(k) = \sharp\{g \in \Gamma \mid \exists g_1, \dots, g_k \in G \text{ such that } g = g_1 \cdot \ldots \cdot g_k\}. \tag{7.2}$$

So $N(k)$ is the number of group elements which can be written as a product of k elements of G. The dependence of $N(k)$ on G is easy to estimate: If G' is another such generating set, then there are numbers p, q such that any element of G can be expressed by p elements of G' and each element of G' by q elements of G. Thus we have

$$N'(k) \geq N(qk), \quad N(k) \geq N'(pk).$$

Theorem 7.7 (*Milnor '68*, [30])
Let M be a complete manifold with $Ric \geq 0$ and let $\Gamma \subset \pi_1(M)$ any finitely generated subgroup of the fundamental group. Then the growth of Γ can be estimated by

$$N(k) < c \cdot k^n. \tag{7.3}$$

where the constant c depends on $\hat{M}$ and the chosen set of generators of Γ.

Proof. Let G be a set of generators as above; it has $N(1)$ elements. Fix a point $o \in \hat{M}$. For all $g \in \Gamma$, let $|g| = |o, go|$. Put $R' = \max\{|g|;\ g \in G\}$. Choose some $r > 0$ small enough, so that

$$B_r(go) \cap B_r(o) = \emptyset \quad \forall g \in \Gamma \setminus \{1\} \tag{7.4}$$

Put $R = R' + r$. Then the family of balls $\{B_r(go);\ g \in G\}$ is disjoint and its union is contained in $B_R(o)$ so that

$$\mathrm{Vol}\big(B_R(o)\big) \geq N(1) \cdot \mathrm{Vol}\big(B_r(o)\big). \tag{7.5}$$

We can iterate this argument as follows: At the second step, we consider

$$G^2 := \{g_1 g_2;\ g_1, g_2 \in G\}.$$

with $\sharp(G^2) = N(2)$. Then all balls $B_r(go)$ with $g \in G^2$ are disjoint and contained in $B_{2R}(o)$ so that

$$\mathrm{Vol}\big(B_{2R}(o)\big) \geq N(2) \cdot \mathrm{Vol}\big(B_r(o)\big). \tag{7.6}$$

In general, we obtain that

$$\mathrm{Vol}\big(B_{kR}(o)\big) \geq N(k) \cdot \mathrm{Vol}\big(B_r(o)\big). \tag{7.7}$$

Recall that we have the Bishop - Gromov inequality (cf. Corollary 5.6),

$$\mathrm{Vol}\big(B_{kR}(o)\big) \leq \omega_n k^n R^n,$$

where ω_n denotes the volume of the euclidean unit ball, hence

$$N(k) \leq \left\{ \frac{\omega_n R^n}{\mathrm{Vol}\big(B_r(o)\big)} \right\} k^n. \tag{7.8}$$

■

8. Gromov's estimate of the Betti numbers

Homology is a main tool to measure the complexity of topology. Fix a field $\mathbf{F}$ and let $H_q(M)$ denote the q-th singular homology of M with coefficients in $\mathbf{F}$. Further, let $H_*(M) = \oplus_{q\geq 0} H_q(M)$ be the total homology of M. The *total Betti number* of M is given by

$$b(M) = \dim_{\mathbf{F}} H_*(M). \tag{8.1}$$

Theorem 8.1 *Gromov, 1980* (cf. [15], [1], [28])
There is a constant $C(n)$ such that:
(a.) any complete n-dimensional manifold M with nonnegative curvature K satisfies

$$b(M) \leq C(n); \tag{8.2}$$

(b.) any compact n-dimensional manifold M with curvature $K \geq -k^2$, and bounded diameter, $diam(M) \leq D$, satisfies

$$b(M) \leq C(n)^{1+kD}. \tag{8.3}$$

We will give the proof of part (a.), following ideas of Abresch [1] and W.Meyer [28]. (Part (b.) is similar, cf. Remark 7.2.) The proof uses the estimates of Bishop-Gromov and Toponogov. It can be viewed as an application of some sort of *Morse theory* for the distance function $\rho(x) = |o, x|$ where $o \in M$ is fixed. In ordinary Morse theory, one considers a smooth function $f : M \to \mathbb{R}$ with isolated critical points with nondegenerate Hessian (p *critical* means that $\nabla f(p) = 0$), and one observes how the topology of $M^c = \{x \in M;\ f(x) < c\}$ is changed as c grows. There are two main facts in Morse theory (cf [29]):
(1.) If $M^b \setminus M^a$ contains no critical points, then M^b and M^a are diffeomorphic.
(2.) If $M^b \setminus M^a$ contains exactly one critical point p, then M^b is homotopic to M^a with a k-cell attached, where k is the index of the Hessian of f at p.

The distance function $\rho = |o,\ | : M \to \mathbb{R}$ is no longer smooth, but we still have the notions of critical and regular points:

Definition 8.2 A point $x \in M$ is called a *regular point* of ρ if there exists $v \in T_xM$ such that

$$\langle v, \gamma'(0)\rangle < 0 \tag{8.4}$$

for any shortest geodesic γ from x to o. Any such vextor v is called *gradientlike*.

A point $x \in M$ is a *critical point* for ρ if it is non-regular, i.e. if for any $v \in T_xM$ there is a shortest geodesic γ from x to o such that

$$\langle v, \gamma'(0)\rangle \geq 0.$$

Remark 8.3 These notions make sense also if the point o is replaced by a closed subset $\Sigma \subset M$. This will be needed in Ch.10.

Fact (1.) is still valid: Since the set of initial vectors of shortest geodesics to o is closed, the gradientlike vectors form an open subset of TM and moreover a convex cone at any regular point. Thus we may cover the closure of $M^b \setminus M^a = B_b(o) \setminus B_a(o)$ by finitely many open sets with gradientlike vector fields and past them together using a partition of unity, thus getting a gradientlike vector field in a neighborhood of the closure of $B_b(o) \setminus B_a(o)$. This has the property that ρ is strictly increasing along its integral curves. Hence, pushing along the integral curves, we may deform the bigger ball $B_b(o)$ into the smaller one $B_a(o)$. (See Lemma 10.9 for details.) We will use this in Lemma 8.10 below.

However, Fact (2.) has no meaning and has to be replaced by another idea: Large balls can be covered by a bounded number of small balls (Bishop-Gromov inequality), and the jump of the Betti number when passing from a small ball to a large ball can be controlled using Toponogov's theorem.

First of all, critical points of ρ are not necessarily isolated, but still in some sense, we have to take only finitely many into account:

Lemma 8.4 Let M be a complete manifold with nonnegative curvature. For any $L > 1$ there exists a finite number $c(n, L)$ such that there are at most $c(n, L)$ critical points $\{q_i\}$ for ρ satisfying

$$|o, q_{i+1}| \geq L|o, q_i|. \tag{8.5}$$

E.g. for $L = 2$ we have $c(n, 2) = 2\sqrt{5}^n$.

Proof. Let $(q_1, q_2, ...)$ be a maximal sequence satisfying (8.5). For $i < j$, let γ be a shortest geodesic from q_i to q_j and put $v = \gamma'(0)$. Since q_i is critical, there is a shortest geodesic c from q_i to o with the angle $\beta = \angle(c'(0), v) \leq 90°$. Applying Toponogov's theorem (Corollary 6.3) with the standard space $\tilde{M} = \mathbb{R}^n$, we get $\tilde{\beta} \leq 90°$.
Consider first the limit case $\tilde{\beta} = 90°$. Let $\tilde{\alpha}$ be the angle in $\tilde{o}$. It follows that

$$\cos(\tilde{\alpha}) = \frac{|q_i, o|}{|q_j, o|} \leq \frac{1}{L}.$$

Fig. 11.

Hence, if $\tilde{\beta} \leq 90°$,

$$\tilde{\alpha}_o \geq \arccos(\frac{1}{L}) =: \alpha_0.$$

Now we apply Toponogov's theorem backwards for the angle at o, but this time we consider an arbitrary shortest geodesic from q_i to q_j. Then for the angle α at o we have

$$\alpha \geq \tilde{\alpha} \geq \alpha_0.$$

It follows as in Ch.7 that there must be a finite number of such critical points. If $L = 2$, we have $\alpha_0 = 60°$ and hence $c(n, 2) = 2\sqrt{5}^n$ as in the proof of Theorem 7.3. ∎

Corollary 8.5 Given a complete manifold M with nonnegative curvature, all critical points are contained in a finite ball.

Since we will work with many metric balls in M, we agree on the following convention: If $B = B_r(p)$ be a fixed ball and $\lambda > 0$, we put $\lambda B := B_{\lambda r}(p)$. More generally, for any $q \in M$ we let $\lambda B(q) := B_{\lambda r}(q)$.

Definition 8.6 Let $A \subset C \subset M$. We define the *content* of A in C as the rank of the inclusion map on the homology level:

$$\operatorname{cont}(A, C) = \operatorname{rk}(i_* : H_*(A) \to H_*(C)) \tag{8.6}$$

Then we define the *content of a metric ball B* as:

$$\operatorname{cont}(B) = \operatorname{cont}(B, 5B). \tag{8.7}$$

Essentially, the content measures the total Betti number of a subset. But it is better than the Betti number since it has a nice monotonicity property: Note that if $A \subset A' \subset B' \subset B \subset M$ then

$$\mathrm{cont}(A, B) \leq \mathrm{cont}(A', B')$$

In the spirit of Morse theory, we will observe how the content of balls grows with the radius. A measure for the number of critical points which are still outside the ball and which will eventually increase the content is the *corank*. This definition involves also critical points of the distance function $\rho_p = |p, \ |$ for points $p \in M$ different from o, called *critical for* p for short.

Definition 8.7 Let $B = B_r(o)$ a ball in M, $r > 0$. For $p \in M$, let $k(r,p)$ be the maximum number of critical points $q_1, \ldots, q_k$ for p such that

1) $|p, q_1| \geq 3Lr$
2) $|p, q_{i+1}| \geq L|p, q_i|$.

By Lemma (8.3), $k(r,p) \leq c(L,n)$. Then we define the *corank* of B as follows:

$$\mathrm{corank}(B) = \inf\{k(r,p) \mid p \in 5B\}. \tag{8.8}$$

Not all balls really contribute to the topology, namely the *compressible* ones:

Definition 8.8 A ball $B \subset M$ is called *compressible* if there is $\tilde{B} = \frac{3}{5}B(q)$ for some $q \in 2B$, and a diffeomorphism $\varphi : M \to M$ such that

$$\varphi|_{M \setminus 5B} = id$$

and

$$\varphi(B) \subset \tilde{B}$$

For short: B is compressible into $\tilde{B}$. Otherwise, the ball B is called *incompressible*.

Lemma 8.9 Suppose that B is compressible into $\tilde{B}$, with $\tilde{B} = \frac{3}{5}B(p)$, and $p \in 2B$. Then

$$\mathrm{cont}(\tilde{B}) \geq \mathrm{cont}(B) \tag{8.9}$$

and

$$\mathrm{corank}(\tilde{B}) \geq \mathrm{corank}(B). \tag{8.10}$$

Proof. Let φ be the diffeomorphism which compresses B into $\tilde{B}$. From $B \approx \varphi(B)$ and

$$\varphi(B) \subset \tilde{B} \subset 5\tilde{B} \subset 5B, \tag{8.11}$$

it follows that $\text{cont}(\tilde{B}) \geq \text{cont}(B)$.

To show the second relation, put $k = \text{corank}(B)$. Let $p \subset 5\tilde{B} \subset 5B$. Then there are $m \geq k$ critical points $q_1, ..., q_m$ for p with

$$|p, q_1| \geq 3Lr \geq 3L \cdot \frac{3}{5}r, \quad |p, q_{i+1}| \geq L|p, q_i|.$$

Hence $k(\frac{3}{5}r, p) \geq m \geq k$ which shows $\text{corank}(\tilde{B}) \geq k$. ∎

Lemma 8.10 If B is incompressible, for any $p \in 2B$ there is a critical point for p in $5\tilde{B} \setminus \tilde{B}$, where $\tilde{B} = \frac{3}{5}B(q)$.

Proof. Otherwise, we could deform $B \subset 5\tilde{B}$ into $\tilde{B}$ while keeping $M \setminus 5B$ fixed, so B would be compressible.

Lemma 8.11 Let B be incompressible, and put $\tilde{B} = \lambda B(\tilde{p})$ for some $\lambda \leq \frac{1}{5L}$ and $\tilde{p} \in \frac{3}{2}B$. Then

$$\text{corank}(\tilde{B}) \geq \text{corank}(B) + 1. \tag{8.12}$$

Proof. Let $\tilde{p} \in \frac{3}{2}B$ and $p \in 5\tilde{B} \subset 2B$. By the previous lemma, there is a critical point q_0 for p in $3B(p) \setminus \frac{3}{5}B(p)$, so we have

$$3r \geq |p, q_0| \geq \frac{3}{5}r \geq 3L\lambda r. \tag{8.13}$$

Now let $k = \text{corank}(B)$. Then there are $m \geq k$ critical points $q_1, ..., q_m$ for p with

$$|p, q_1| \geq 3Lr, \quad |p, q_{i+1}| \geq L|p, q_i|. \tag{8.14}$$

Thus by (8.13),

$$|p, q_1| \geq L|p, q_0|, \quad |p, q_0| \geq 3L\lambda r. \tag{8.15}$$

Now (8.14) and (8.15) show

$$k(\lambda r, p) \geq m + 1 \geq k + 1.$$

Since $p \in 5\tilde{B}$ was arbitrary, we get

$$\text{corank}(\tilde{B}) \geq k + 1.$$

∎

Proof of theorem 8.1

Let $K = \max\{\text{corank}(B) \mid B \text{ ball in } M\}$. By Lemma 8.4, $K \leq c(n, L)$, and by Corollary 8.5, for a large enough ball B , there exists an isotopy of M that carries M into B, hence

$$\text{cont}(B) = \text{cont}(M, M) = b(M). \tag{8.16}$$

Gromov's theorem now follows if we prove that the content of any metric ball is bounded by a constant. This is done in 5 steps:

Step 1. If $\text{corank}(B) = K$ then $\text{cont}(B) = 1$.

Proof: Let $p_0 := p$. By Lemma 8.11, B is compressible into $B_1 = \frac{3}{5}B(p_1)$ for some $p_1 \in 2B$ (otherwise, we could increase the corank). Hence, $|p_0, p_1| < 2r$. By Lemma 8.9, $\text{corank}(B_1) = K$ and $\text{cont}(B_1) \geq \text{cont}(B)$. Repeating the argument, we compress B_1 (and hence B) into $B_2 = \frac{3}{5}B_1(p_2) = (\frac{3}{5})^2 B(p_2)$ for some $p_2 \in 2B_1$. Hence $|p_1, p_2| < \frac{3}{5} \cdot 2r$. After N steps, we have compressed B into $B_N = (\frac{3}{5})^N B(p_N)$ with $|p_i, p_{i+1}| < (\frac{3}{5})^i \cdot 2r$ for $i = 0, ..., N$. So we have for all $q \in B_N$:

$$|p, q| < \sum_{i=0}^{N} (\frac{3}{5})^i \cdot 2r < 5r.$$

Thus $B_N \subset 5B$. Since the cut locus distance (injectivity radius) is bounded below on the closure of B (by compactness), we find some N such that $(\frac{3}{5})^N r$ is smaller than this bound which implies that B_N is diffeormorphic to a euclidean ball and hence contractible. So, $\text{cont}(B_N) = 1$ since $b(B_N) = 1$. On the other hand,

$$\text{cont}(B_N) \geq \text{cont}(B_{N-1}) \geq ... \geq \text{cont}(B)$$

which proves $\text{cont}(B) = 1$.

Step 2. Each ball B with $\text{cont}(B) \geq 2$ is either incompressible or it contains an incompressible ball $\tilde{B}$ with $\text{cont}(\tilde{B}) \geq \text{cont}(B)$ and $\text{corank}(\tilde{B}) \geq \text{corank}(B)$.

Proof. Otherwise, we could use the process of Step 1 to show that $\text{cont}(B) = 1$.

Step 3. If B is incompressible, then any ball $\tilde{B} = \lambda B(p)$ with $\lambda \leq \frac{1}{5L}$ and $p \in \frac{3}{2}B$ satisfies

$$\text{corank}(\tilde{B}) \geq \text{corank}(B) + 1. \tag{8.17}$$

Proof. Cf. Lemma 8.11.

Step 4. There is a number $\chi \leq (2L)^{n(n+1)} 10^{n(n+1)(n+2)}$ with the following property: Suppose that any ball $\tilde{B}$ with $\text{corank}(\tilde{B}) \geq k$ has $\text{cont}(\tilde{B}) \leq a_k$. Then for any ball B with $\text{corank}(B) \geq k - 1$ we have

$$\text{cont}(B) \leq \chi \cdot a_k \tag{8.18}$$

Proof. By Step 2 we may assume that B is incompressible. Let $\lambda = 1/(5L \cdot 10^{n+1})$. We cover B with balls $B_i = \lambda B(p_i)$ with $p_i \in B$ for $i = 1, ..., N$ such that the balls $\frac{1}{2}B_i$ are disjoint and inside B. Let $\frac{1}{2}B_1 = \frac{1}{2}\lambda B(p_1)$ the one of smallest volume among $\frac{1}{2}B_1, ..., \frac{1}{2}B_N$. Since $B \subset 2B(p_1)$, we have

$$\mathrm{vol}(2B(p_1)) \geq \mathrm{vol}(B) \geq \sum_{i=1}^{N} \mathrm{vol}(\frac{1}{2}B_1) \geq N \cdot \mathrm{vol}(\frac{1}{2}B_1) = N \cdot \mathrm{vol}(\frac{1}{2}\lambda B(p_1)).$$

On the other hand, by Bishop-Gromov (cf. Corollary 5.6),

$$\frac{\mathrm{vol}(2B(p_1))}{\mathrm{vol}(\frac{1}{2}\lambda B(p_1)}) \leq \frac{\omega_n \cdot (2r)^n}{\omega_n \cdot (\frac{1}{2}\lambda r)^n} = (\frac{4}{\lambda})^n,$$

thus

$$N \leq (\frac{4}{\lambda})^n = (2L \cdot 10^{n+2})^n. \tag{8.19}$$

By Step 3, all $B_i^m := 10^m B_i$ for $m = 0, ..., n+1$ have corank $\geq k$ and thus $\mathrm{cont}(B_i) \leq a_k$. Since the radii of the $B_i = B_i^0$ are very small, we have (by triangle inequality)

$$B \subset \bigcup_i B_i \subset \bigcup_i 10^{n+1} B_i \subset 5B,$$

hence

$$\mathrm{cont}(B) \leq \mathrm{cont}(\bigcup_i B_i, \bigcup_i 10^{n+1} B_i).$$

We will see below (cf. Appendix) that we may estimate

$$\mathrm{cont}(\bigcup_i B_i, \bigcup_i 10^{n+1} B_i) \leq \sum_{j=1}^{n+1} \sum_{i_1 < ... < i_j} \mathrm{cont}(\bigcap_{p=1}^{j} B_{i_p}^{n-j+1}, \bigcap_{p=1}^{j} 10 B_{i_p}^{n-j+1}) \tag{8.20}$$

where $B_i^m := 10^m B_i$. Since

$$\bigcap_{p=1}^{j} B_{i_p}^m \subset B_{i_1}^m \subset 5B_{i_1}^m \subset \bigcap_{p=1}^{j} 10 B_{i_p}^m,$$

we get

$$\mathrm{cont}(\bigcap_{p=1}^{j} B_{i_p}^m, \bigcap_{p=1}^{j} 10 B_{i_p}^m) \leq \mathrm{cont}(B_{i_1}^m) \leq a_k.$$

There are N balls $B_1^m, ..., B_N^m$, so there are at most N^n intersections $B_{i_1}^m \cap ... \cap B_{i_j}^m$ where $j = 1, ...n$. Thus

$$\text{cont}(B) \leq N^{n+1} a_k$$

which finishes the proof by (8.19)

Step 5. Let a_k still denote an upper bound for the content of any ball with corank $\geq k$. By Step 1 we may choose $a_K = 1$ where $K \leq c := c(L, n)$ is the biggest possible corank. Thus by Step 4, $a_{K-1} = \chi$, hence $a_{K-2} = \chi^2$ and eventually (by induction), $a_0 = \chi^K$. Hence we get for any ball B

$$\text{cont}(B) \leq \chi^c$$

which finishes the proof since $b(M) = \text{cont}(B)$ for some big ball (cf. (8.16)). ■

Remark 8.12 The highest known total Betti number for a manifold with $K \geq 0$ is 2^n, the total Betti number of the n-dimensional torus $T^n = S^1 \times ... \times S^1$.

9. Convexity.

Definition 9.1 Let M be a Riemannian manifold. A continuous function $f : M \to \mathbb{R}$ is called *convex* if $f \circ \gamma : I \to \mathbb{R}$ is convex for any geodesic $\gamma : I \to M$, and f is called *concave* if $-f$ is convex.

Theorem 9.2 A continuous function $f : M \to \mathbb{R}$ is convex if for any $x \in M$ and $\varepsilon > 0$ there is a smooth *lower support function* $f_{x,\varepsilon} = \tilde{f} : U_x \to \mathbb{R}$ (i.e., $\tilde{f} \le f$, $\tilde{f}(x) = f(x)$), defined in a neighborhood $U_x \subset M$, with $D\nabla\tilde{f}(x) > -\varepsilon$.

Proof. Suppose there is a geodesic $\gamma : I \to M$ such that $g = f \circ \gamma$ is not convex. Then there exists a parabola $a : \mathbb{R} \to \mathbb{R}$ with:
1) $a'' \equiv -\varepsilon$
2) $a(t_1) = g(t_1)$ and $a(t_2) = g(t_2)$
3) there exists $t \in (t_1, t_2)$ with $a(t) < g(t)$.

Fig. 12.

Define $\delta = g - a$; then δ takes a maximum at some point $t_0 \in (t_1, t_2)$, and we have $\delta(t_0) > 0$. Consider $\tilde{f} = f_{x,\varepsilon}$ with $x = \gamma(t_0)$ and ε small enough. Then $\tilde{g} = \tilde{f} \circ \gamma$ is a lower support function for g at t_0, hence $(\tilde{g} - a)$ also takes a maximum at t_0. But $g''(t_0) > -\varepsilon$ and $a''(t_0) = -\varepsilon$, hence

$$(\tilde{g} - a)''(t_0) > -\varepsilon + \varepsilon = 0$$

which is impossible at a maximum point. ∎

Remark 9.3 By the theorem, a C^2-function f is convex if $D\nabla f \ge 0$. If f is only continuous but satisfies the assumptions of Theorem 9.2, we say $D\nabla f \ge 0$ *in the sense of support functions.*

Example 9.4 Fix $o \in M$ and let $\rho(x) = |o, x|$. Then $D\nabla\rho = A(\rho) > 0$ for small $\rho > 0$ by example 2.3. However, ρ is not smooth at o, but ρ^2 is smooth (since $\rho(\exp_o(v))^2 = \langle v, v \rangle$ if $\|v\|$ is small, cf. (6.1)), and it satisfies

$$D\nabla(\rho^2) = 2(d\rho \cdot \nabla\rho + \rho \cdot D\nabla\rho) > 0,$$

hence ρ^2 is convex near o.

Definition 9.5 A closed subset $C \subset M$ is called *convex* if any geodesic segment γ in M with end points on C lies entirely in C. Clearly, if $f : M \to \mathbb{R}$ is a convex function, then the sublevel sets $M^a = \{x \in M \mid f(x) \leq a\}$ are convex subsets for all $a \in \mathbb{R}$. Note that the notion of convexity depends on the surrounding manifold. For example, in the cylinder $S^1 \times \mathbb{R}$ with radius 1, a metric ball of radius $r < \pi$ is not convex, but it is convex in a slightly larger ball of radius $r + \varepsilon < \pi$ inside the cylinder.

Definition 9.6 A smooth hypersurface $S = \partial B \subset M$ is called *convex hypersurface* if for the interior unit normal vector field N,

$$DN \leq 0. \tag{9.1}$$

Clearly, if f is smooth and convex, then $S = \partial M^a$ is a convex hypersurface unless a is a minimum of f (note that ∇f can vanish only at a minimum of a convex function f). Vice versa, if $S = \partial B$ is a *strictly* convex hypersurface, i.e. $DN < 0$, then the signed distance function $\rho = \rho_S = \pm|\ , S|$ is concave near S since DN is its Hessian along S (cf. Ch.2, p.9). Consequently, $C = Clos(B)$ is convex in a neighborhood of C since it is a sublevel set of the convex function $-\rho$.

Theorem 9.7 Let B be compact and $S = \partial B$ a convex hypersurface. If $K \geq 0$ on B, then $\rho = |\ , S|$ is concave on all of B.

Proof. Let $x \in B$ and γ a shortest geodesic from x to S. Let $p \in S$ be its endpoint. If we were in euclidean space, there would be a support hyperplane for B at p. In M we have a similar construction: For large R, let $\tilde{S} = \exp_p(\partial B_R(-RN_p) \cap U)$ where U is a small neighborhood of $-RN_p$ in T_pM. We will show that $\tilde{S}$ supports B at p, i.e. $\tilde{S} \cap B = \emptyset$ and $\tilde{S} \cap \partial B = \{p\}$. To this end, we compare the signed distance functions ρ of S and $\tilde{\rho}$ of $\tilde{S}$ near p. Since N_p is a common normal vector for S and $\tilde{S}$, these functions agree at p up to first derivatives. The second derivatives (Hessian) of ρ and $\tilde{\rho}$ are given by the shape operators $A = (DN)_p$ and $\tilde{A} = (D\tilde{N})_p$ of S and $\tilde{S}$. By convexity, we have $A \leq 0$. On the other hand, $\tilde{A} = \frac{1}{R} \cdot I > 0$. (In fact, this holds for the euclidean sphere $\partial B_R(-RN_p) \subset T_pM$ and hence also for $\tilde{S}$ since $\exp_p$ preserves the covariant derivative

D at p, cf. Remark 1.5.) Thus $A < \tilde{A}$. However, the Hessians of ρ and $\tilde{\rho}$ are both zero in N_p-direction, so we have no strict inequality. This can be repaired by passing to functions $f \circ \rho$ and $g \circ \tilde{\rho}$ which have the same level hypersurfaces as ρ and $\tilde{\rho}$, where we choose the functions $f, g : \mathbb{R} \to \mathbb{R}$ monotone with

$$f(0) = g(0), \;\; f'(0) = g'(0), \;\; f''(0) < g''(0).$$

Then $f \circ \rho$ and $g \circ \tilde{\rho}$ agree at p up to first derivatives with $Dd(f \circ \rho)_p < Dd(g \circ \tilde{\rho})_p$, so we get $f \circ \rho \leq g \circ \tilde{\rho}$ near p (in fact, "$<$" outside p). In particular, the level set $\tilde{S} = \{g \circ \tilde{\rho} = 0\}$ lies outside $B = \{f \circ \rho > 0\}$.

Fig. 13.

Now we consider the signed distance function $\tilde{\rho}$ of $\tilde{S}$ on its full domain, namely wherever the mapping

$$S \times \mathbb{R} \to M : \;\; (s, t) \mapsto \exp(tN(s))$$

is invertible, cf. Ch.2, p.8. In particular, ρ is defined and smooth near x: Since γ is a shortest curve from p to S, there are no focal points for S on γ between p and x. Since $\tilde{A} > A$ at p, the focal distance for $\tilde{S}$ along γ is strictly larger than for S (cf. Remark 3.2), hence x is no focal point for $\tilde{S}$, and $\tilde{\rho}$ is defined and smooth near x.

Since $\tilde{S}$ lies outside B, any curve from B to $\tilde{S}$ meets $S = \partial B$ first, so $\tilde{\rho}$ is an upper support function for ρ at x. Moreover, by the comparison theorem 3.1 we have

$$D\nabla\tilde{\rho}(x) \leq \frac{1}{R + \tilde{\rho}(x)} =: \varepsilon,$$

so $-\rho_S$ is convex on B by theorem 9.1. ■

Remark 9.8 The proof shows that $S = \partial B$ need not to be smooth; it is sufficient that there is a unit vector $N_p \in T_pM$ such that $\tilde{S}_{p,R} := \exp_p(\partial B_R(-RN_p) \cap U)$ supports B for any $p \in S$, $R > 0$ and a neighborhood U of $-RN_p$ in T_pM. For example, this is true if $Clos(B) =: C$ is convex (with $B = Int(C)$). To see this, let $p \in \partial C$ and consider the *tangent cone*

$$T_pC = Clos\{v \in T_pM;\ \exp_p(tv) \in C \text{ for small } t > 0\}.$$

By convexity of C, this is a convex cone in T_pM (since in exponential coordinates, D-geodesics and ∂-geodesics, i.e. straight lines, are close to each other near the origin, see 1.5), hence it is contained in some closed half space $H_N = \{v \in T_pM;\ \langle v, N\rangle \geq 0\}$. Any such vector $N = N_p$ is called an *inner normal vector* of ∂C at p, and $-N_p$ is called an *outer normal vector*. By strict convexity of $\tilde{S} = \tilde{S}_{p,R}$, the set $\tilde{D} = \exp_p(D_R(-RN_p) \cap U)$ is convex in a neighborhood of p, being a sublevel set of the convex function $-\tilde{\rho}$. Thus, $\tilde{D} \cap C$ is convex near p. But this shows that $\tilde{D} \cap C = \{p\}$, proving that $\tilde{S}$ supports C at p. Namely, if $q = \exp_p v \in \tilde{D} \cap C$, then $\exp_p tv \in \tilde{D} \cap C$ for all $t \in [0, 1]$, hence $v \in T_pC \cap D_R(-RN_p) = \{0\}$, so $q = p$.

10. Open manifolds with nonnegative curvature

Theorem 10.1 (*Cheeger - Gromoll 1970* [6])
Let M an *open* (i.e. complete, noncompact) manifold with $K \geq 0$. There exists a compact convex submanifold (without boundary) $\Sigma \subset M$ called *soul* such that M is diffeomorphic to the normal bundle $\nu\Sigma$.

Recall that the *normal bundle* of a submanifold $\Sigma \subset M$ is $\nu\Sigma = \cup_p \nu_p\Sigma$, where $\nu_p\Sigma = \{v \in T_pM \mid v \perp T_p\Sigma\}$. The proof needs some preparations. For the moment, assume only that M is complete and noncompact.

Definition 10.2 A *ray* is a geodesic γ defined on $[0, \infty)$ which is a shortest geodesic between any two of its points.

Remark 10.3 For any $p \in M$ there exists a ray starting at p: Consider a sequence of points q_i such that $|q_i, p| \to \infty$. Consider shortest geodesics γ_i from p to q_i. There exists a subsequence of the unit tangent vectors converging to some $v \in S_pM$. The geodesic γ_v in the direction of v is a ray. Moreover, if the points q_i lie on some geodesic ray γ, i.e. $q_i = \gamma(t_i)$ with $t_i \to \infty$, the geodesic ray γ_v is called an *asymptote* of γ. Note that asymptotes are not necessarily unique.

Definition 10.4 The *Busemann function* associated to a ray γ is defined as

$$b_\gamma(x) = \lim_{t\to\infty} (|x, \gamma(t)| - t) \tag{10.1}$$

In particular, $b_\gamma(\gamma(s)) = \lim(|\gamma(s), \gamma(t)| - t) = \lim(t - s - t) = -s$. Further, since $\rho_{\gamma(t)} = |\cdot, \gamma(t)|$ are Lipschitz-continuous functions with Lipschitz constant $L = 1$, the same holds for $b_\gamma(x)$. Its level sets are called *horospheres*.

Consider an asymptote γ_x of the ray γ. We define

$$b_{x,t}(y) = |\gamma_x(t), y| - t + b_\gamma(x). \tag{10.2}$$

$b_{x,t}$ is smooth in a neighborhood of p since x is not in the cut locus of any point on γ_x.

Lemma 10.5 $b_{x,t}$ is a support function of b_γ at x.

Proof. There is a sequence $t_i \to \infty$ such that $\gamma_x = \lim \gamma_i$ where γ_i is a shortest geodesic from x to $\gamma(t_i)$. Then by triangle inequality, we have for any $y \in M$:

$$\begin{aligned} b_{x,t}(y) &= |\gamma_x(t), y| - t + b_\gamma(x) \\ &\approx |\gamma_i(t), y| - t + b_\gamma(x) \\ &\geq |\gamma(t_i), y| - s_i + b_\gamma(x) \end{aligned}$$

Fig. 14.

where $s_i = |x, \gamma(t_i)|$. The sign "$\approx$" means that the error can be made as small as one wants (as $i \to \infty$). From

$$b_\gamma(x) \approx |x, \gamma(t_i)| - t_i = s_i - t_i$$

we get

$$\begin{aligned} b_{x,t}(y) &\geq |y, \gamma(t_i)| - t_i \\ &\approx b_\gamma(y). \end{aligned}$$

Moreover, $b_{x,t}(x) = b_\gamma(x)$. So $b_{x,t}$ is a support function to b_γ. ■

Lemma 10.6 If M is complete, noncompact with $K \geq 0$, then each Busemann function $b_\gamma(x)$ is concave.

Proof. It follows from the comparison theorem 3.1 with $k = 0$ that

$$D\nabla b_{x,t}(x) \leq \frac{1}{|x, \gamma_x(t)|} \to 0$$

as $t \to \infty$. So by Theorem 9.2, b_γ is concave. ■

Corollary 10.7 The superlevel sets

$$C_{t,\gamma} = \{x \in M \mid b_\gamma(x) \geq t\} \tag{10.3}$$

are convex sets in M for any $t \in \mathbb{R}$.

For any point $p \in M$ we consider

$$R_p = \{\text{rays } \gamma : [0, \infty) \to M, \quad \gamma(0) = p\}. \tag{10.4}$$

Define the function

$$b = \min_{\gamma \in R_p} b_\gamma. \tag{10.5}$$

(The infimum is a minimum since a limit of rays is a ray.) b is concave since it is the minimum of concave functions, and therefore, its superlevel sets

$$C_t = \{x \in M \mid b(x) \geq t\}. \tag{10.6}$$

are convex.

Lemma 10.8 C_t is compact.

Proof. Assume first $t \leq 0$. If C_t were not compact, there whould exist a sequence of points $q_i \to \infty$ in C_t. Since $b(p) = 0$, we have $p \in C_t$. For any i, choose a shortest geodesic segment γ_i from p to q_i; since C_t is convex, γ_i is contained in C_t. Since C_t is closed, the limit ray $\gamma = \lim_i \gamma_i$ again lies in C_t. But γ is a ray starting at p, so $\gamma \in R_p$. But then γ cannot be contained in C_t since $b(\gamma(s)) \leq b_\gamma(\gamma(s)) = -s$ can be made smaller than t, contradiction!
Since $C_s \subset C_t$ for $s > t$, all superlevel sets of $b(x)$ are compact. ■

Now let t_0 be the maximum value of b (which exists by compactness). Define $C^1 = C_{t_0}$. C^1 cannot contain interior points. (In fact, if $x \in C^1$ and $b(x) = b_\gamma(x)$, then b_γ decreases with speed one along an asymptotic ray γ_x of γ. Thus $b \leq b_\gamma$ cannot stay maximal near x.) Thus C^1 is a compact convex set of lower dimension. In general, a compact convex subset C of a Riemannian manifold is a subset of a (non complete) submanifold M^1, such that C has nonempty interior relative to M^1 (cf. [5]). If $\partial C^1 \neq \emptyset$, we consider the distance function $\rho_{\partial C_1}$ on C^1. Since C^1 is convex, ∂C^1 is a convex hypersurface. ¿From Theorem 9.7, and Remark 9.8 we see that $\rho_{\partial C_1}$ is concave on C^1, thus its superlevel sets

$$C^1_t = \{\rho_{\partial C_1} \geq t\}$$

are convex again. If t_1 is the maximum of $\rho_{\partial C_1}$, the set

$$C^2 = C^1_{t_1}$$

cannot have interior points in M^1, thus is again of lower dimension. In this way we produce a descending chain

$$C^1 \supset C^2 \supset \ldots \supset C^k$$

of compact convex subsets with lower and lower dimension. This process ends after a finite number (say: k) of steps; if we put $\Sigma = C^k$, then Σ is a compact convex set without boundary: the soul of M.

It remains to show that M is diffeomorphic to $\nu\Sigma$. In fact, we will show that M is diffeomorphic to a tubular neighborhood of Σ, say $B_r(\Sigma)$, for small r which itself is diffeomorphic to $\nu\Sigma$ via the exponential map $\exp|\nu\Sigma$.

Lemma 10.9 For small $r > 0$, there is a diffeomorphism $\varphi : M \to B_{2r}(\Sigma)$ with $\varphi = id$ on $B_r(\Sigma)$.

Proof. Let $\rho_\Sigma = |\ ,\Sigma|$. We show first that ρ_Σ has no critical points on $M \setminus \Sigma$, i.e. for any $x \in M \setminus \Sigma$ there is a gradientlike vector $v \in T_xM$ for ρ_Σ.
In fact, since $x \notin \Sigma$, there is $j \in \{1, \ldots, k-1\}$ such that $x \in C^j \setminus C^{j+1}$. In particular, there is some t such that $x \in \partial C^j_t$. Now, since $\Sigma \subset C^j_t$, any geodesic γ from x to Σ has initial vector $\gamma'(0)$ pointing to the interior of C^j_t. Thus, an outer normal vector for the convex set C^j_t is gradientlike.

Now for any $x \in M \setminus B_r(\Sigma)$, we choose a gradientlike vector $v \in N_xC_t$ and enlarge it to a gradientlike vector field V_x on some neighborhood U_x. By paracompactness, we may pass to a locally finite subcovering $(U_{x_i})_{i=1,2,\ldots}$. Let $V_i = V_{x_i}$. On $B_{2r}(\Sigma) \setminus \Sigma$, we put $V_0 = \nabla\rho_\Sigma$: We have chosen $r > 0$ so small that ρ_Σ is smooth on $B_{2r}(\Sigma)$; this is possible by compactness of Σ. We may choose a subordinated partition of unity $(\varphi_i)_{i\geq 0}$: then $V = \sum \varphi_i V_i$ is a smooth gradientlike vector field defined on $M \setminus \Sigma$ which agrees to $\nabla\rho_\Sigma$ on $B_r(\Sigma) \setminus \Sigma$.

Let c_x be the integral curve of V with $c_x(0) = x$. Since the integral curves intersect $\partial B_r(\Sigma)$ transversally (in fact, orthogonally), there is a smooth function

$$t : M \setminus \Sigma \to \mathbb{R}$$

with

$$c_x(t(x)) \in \partial B_r(\Sigma). \tag{10.7}$$

We reparametrize the integral curves and put

$$\tilde{c}_x(t) = c_x(r + t(x) - t).$$

Now let $\chi : \mathbb{R}_+ \to [0, 2r)$ smooth, with $\chi(t) = t$ for $t \in [0, r]$ and $\chi' > 0$,
and let $\varphi : M \to B_{2r}(\Sigma)$ with

$$\varphi(x) = \begin{cases} x & x \in B_r(\Sigma) \\ \tilde{c}_x(\chi(r + t(x))) & \text{otherwise} \end{cases}$$

Then φ maps diffeomorphically M onto $B_{2r}(\Sigma)$. ■

Fig. 15.

Theorem 10.10 (*Perelman*) Let M be an open manifold of $K \geq 0$ with soul Σ. For any $p \in \Sigma$ and any two nonzero vectors $a \in T_p\Sigma$, $v \in \nu_p\Sigma$, there is a totally geodesic flat half plane through p tangent to a and v.

The proof uses the contraction of Sharafutdinov ([32],[35]) which is a continuous mapping $\phi : M \to \Sigma$ with

$$|\phi(x), \phi(y)| \leq |x, y|$$

for all $x, y \in M$. In fact, ϕ is the limit of an iterated projection onto more and more smaller and smaller convex sets surrounding Σ, cf. [35] for details.

Proof of Perelman's theorem: Let

$$F = \phi \circ exp : \nu\Sigma \to \Sigma$$

and put

$$f(v) = |\pi(v), Fv|$$

for $v \in \nu\Sigma$, where $\pi : \nu\Sigma \to \Sigma$ denotes the projection. Consider

$$mf(t) := \max\{f(v) \; ; \quad v \in \nu^t\Sigma\}$$

where $\nu^t\Sigma = \{v \in \nu\Sigma; \; \|v\| = t\}$. Clearly $mf(0) = 0$ and $mf(t) \geq 0$ for all $t \geq 0$. We claim that $mf(t)$ is monotonely decreasing so that we get $mf \equiv 0$. Let $t > 0$ be small enough such that $mf(t)$ is strictly less than the cut locus distance on Σ. Let $v \in \nu^t\Sigma$ so that $f(v) = mf(t)$. Let α be the shortest geodesic in Σ from $F(v)$ to $p = \pi(v)$. By assumption, the prolongation of α stays shortest beyond p up to some point $q \in \Sigma$. Let $w \in \nu_q\Sigma$ be the parallel displacement of v from p to q along α.

Fig. 16.

Lemma 10.11 We have $f(w) = mf(t)$ and $|Fv, Fw| = |p, q|$, and $p, q, \exp v, \exp w$ span a totally geodesic flat rectangle R.

Proof. By the contraction property of ϕ and Rauch II we have

$$|Fv, Fw| \leq |\exp v, \exp w| \leq |p, q|, \tag{10.8}$$

and moreover, by the maximality of $f(v) = mf(t)$,

$$|Fw, q| = f(w) \leq mf(t). \tag{10.9}$$

On the other hand, Fv, p and q are lined up on a shortest geodesic, so

$$|Fv, q| = mf(t) + |p, q|. \tag{10.10}$$

So the triangle inequality for Fv, Fw, q yields equality in (10.9) and (10.10) and we have proved the Lemma, using the equality case of Rauch II (cf. Corollary 3.4). ∎

Now we consider the vector $w' = (1 - (\varepsilon/t))w$ with length $\|w'\| = \|w\| - \varepsilon = t - \varepsilon$. Clearly

$$mf(t - \varepsilon) \geq f(w') = |q, Fw'|.$$

Since $R = (p, \exp v, \exp w, q)$ is a flat rectangle,

$$|\exp v, \exp w'|^2 \leq |\exp v, \exp w|^2 + \varepsilon^2 = |p, q|^2 (1 + (\frac{\varepsilon}{|p, q|})^2),$$

hence

$$|Fv, Fw'| \leq |\exp v, \exp w'| \leq |p, q| + C\varepsilon^2 \tag{10.11}$$

Fig. 17.

with $C = (2|p,q|)^{-1}$. Using (10.10),(10.11) and the triangle inequality for Fv, q, Fw' we get

$$mf(t-\varepsilon) \geq |q, Fw'| \geq |Fv, q| - |Fv, Fw'| \geq mf(t) - C\varepsilon^2.$$

This implies that mf is monotonely decreasing and hence zero. Thus $f \equiv 0$ and $\phi(\exp v) = \pi(v)$ for any $v \in \nu\Sigma$. (Note that the condition for $t = \|v\|$ made at the beginning now is void.) In particular, $|\exp v, \exp w| = |\pi v, \pi w|$ for any two normal vectors v, w on Σ which are parallel along a shortest geodesic on Σ. So the proof is finished by the equality case of Rauch II (Corollary 3.4). ∎

Remark 10.12 The above theorem also proves a conjecture of Cheeger and Gromoll saying that the soul must be a point (and hence M must be diffeomorphic to $\mathbb{R}^n$) provided that there is a point $q \in M$ where all sectional curvatures are strictly positive. In fact, we can connect q to Σ by a shortest geodesic γ (which is perpendicular to Σ). If Σ has positive dimension, i.e. if there is a nonzero tangent vector a of the soul where it meets γ, then by Perelman's theorem, there is a flat totally geodesic half plane spanned by a and γ. Thus not all curvatures at $q = \gamma(0)$ are positive. - We are indepted to V. Schroeder for communicating Perelman's proof.

Remark 10.13 By a previous result of Strake and Walshap [33], Perelman's theorem implies that on a small tubular neighborhood of the soul, the projection $\pi : B_r(\Sigma) \to \Sigma$ is a Riemannian submersion, i.e. $d\pi_x$ is an orthogonal projection up to isometric linear isomorphisms, for any $x \in B_r(\Sigma)$. We conjecture that the Sharafutdinov contraction $\phi : M \to \Sigma$ is smooth and also a Riemannian submersion.

11. The sphere theorem.

One of the most celebrated results in Riemannian geometry is the "sphere theorem"; cf. [2], [5], [11], [18], [25], [28].

Theorem 11.1 (*Rauch, Berger, Klingenberg*)
Let M^n be a compact, simply connected manifold, with $K > 0$. Assume that

$$\frac{\max K}{\min K} < 4. \tag{11.1}$$

Then M is homeomorphic to S^n.

Remark 11.2 The estimate 4 is sharp. There are simply connected manifolds with

$$\frac{\max K}{\min K} = 4$$

which are not homeomorphic to S^n, namely the projective spaces over the fields $\mathbb{C}$ and $\mathbb{H}$ (called $\mathbb{C}P^m$ and $\mathbb{H}P^m$), and the Cayley projective plane.

Remark 11.3 Another type of sphere theorem using a diameter estimate instead of the upper curvature bound was given by Grove and Shiohama ([21], [28]).

Proof. We may assume (rescaling the metric) that

$$\frac{1}{4} < K < 1 \tag{11.2}$$

so the comparison spaces are the spheres S_2 and S_1 with radii 2 and 1 respectively. We fix $p \in M$, and consider geodesics of length π on the three manifolds.

Fig. 18.

By the upper curvature bound, the first conjugate point comes later than on S_1, hence the exponential map $\exp_p|\overline{B_\pi(0)}$ is an immersion (local diffeomorphism).
By the lower curvature bound, the immersed hypersurface $f := \exp_p|\partial B_\pi(0)$ is strictly concave, i.e. $DN < 0$ for the exterior unit normal field N (cf. Theorem 3.1)

Now we need Theorem 11.4 (see below) to finish the proof. Take two copies D_+, D_- of the unit disk D^n and identify D_+ with $D_\pi(0) = \overline{B_\pi(0)}$ and put $S = \partial B_\pi(0)$. By Theorem 11.4, there exists an immersion $F : D_- \to M$ and a local diffeomorphism $\varphi : S = \partial D_+ \to \partial D_-$ such that $f = F \circ \varphi$. Let

$$\Sigma_\varphi = D_+ \cup_\varphi D_- = (D_+ \amalg D_-)/\sim, \tag{11.3}$$

where the equivalence relation is given as follows: $x \in \partial D_+$ and $y \in \partial D_-$ are equivalent $(x \sim y)$ iff $y = \varphi(x)$.

Fig. 19.

This is a smooth manifold which is homeomorphic to S^n (see figure). It is diffeomorphic if and only if φ extends to a diffeomorphism $\phi : D_+ \to D_-$. Now we define $\hat{F} : \Sigma_\varphi \to M$ by putting

$$\hat{F}|D_+ = \exp_p\,, \quad \hat{F}|D_- = F. \tag{11.4}$$

Then $\hat{F}$ is a local diffeomorphism. It is onto because the image is open and closed and it is injective, since M is simply connected. Hence $\hat{F}$ is a diffeomorphism. ∎

Remark 11.4 In dimension $n = 2$ we cannot apply the above proof. However, the theorem follows from the Gauss-Bonnet formula: Since $K > 0$, it follows that

$$0 < \int_M K = 2\pi\chi(M), \tag{11.5}$$

thus $\chi(M) = (2 - 2g) > 0$, which implies that $g = 0$ where g is the genus of M. Hence M is a sphere.

Theorem 11.5 (*Gromov, Eschenburg* [11])
Let M^n, $n \geq 3$, be a complete manifold with $K \geq 0$. Let S be a compact $(n-1)$-dimensional manifold, and $f : S \to M$ an ε-*convex* immersion, i.e. we suppose that a unit normal vector field N along f satisfies $DN < -\varepsilon$. Then f bounds an immersed convex disk, i.e. there exists an immersion

$$F : D^n \to M \tag{11.6}$$

(where D^n is the unit n-disk) and a diffeomophism

$$\varphi : S \to \partial D^n \tag{11.7}$$

such that $f = \varphi \circ F$ and $N \circ \varphi^{-1}$ becomes the interior normal field.

If $M = \mathbb{R}^n$, we get the following stronger theorem which is due to Hadamard [22]:

Theorem 11.6 (*Hadamard*)
Let S be a compact $(n-1)$-manifold, and $f : S \to \mathbb{R}^n$ an ε-convex immersion. Then f is an embedding and $f(S)$ bounds a convex n-disk.

Proof. Let N be the unit normal field, considered as *Gauss map* $N : S \to S^{n-1}$. Since S is strictly convex, dN has only positive eigenvalues (note that $DN = dN$ where D is the Levi-Civita derivative on $\mathbb{R}^n$, i.e the ordinary derivative $D = \partial$), so N is a local diffeomorphism. But S^{n-1} is simply connected for $n \geq 3$, hence N is a diffeomorphism. So any vector $v \in S^{n-1}$ arises exactly twice as a normal vector for f, namely $v = N(x) = -N(y)$ for exactly two points $x, y \in S$. Thus, any height function $\langle f, v\rangle$ has exactly two critical points: one maximum and one minimum. From this we see that f is injective: Since the height function $\langle f, v\rangle$ for $v = N(x)$ attains its maximum at x, there is no other point $y \in S$ with $\langle f(y), v\rangle = \langle f(x), v\rangle$ (otherwise there would be a second maximum). Moreover, we see that $f(S)$ bounds the convex set

$$C = \bigcap_{x \in S} H(x)$$

where $H(x)$ is the half space

$$H(x) = f(x) + \{v \in \mathbb{R}^n \; ; \; \langle v, N(x)\rangle \leq 0\}.$$

■

Remark 11.7 If $M \neq \mathbb{R}^n$, it is possible to construct an $\varepsilon-$convex immersion that is not injective. For example, take the cylinder $S^1 \times \mathbb{R}$ with radius 1, and consider the immersion $\exp_p |\partial B_r(0)$ for $r > \pi$ (for arbitrary $p \in M$).

Fig. 20.

Remark 11.8 Theorems 11.4 and 11.5 are false for $n = 2$: a counterexample is given by any locally strongly convex closed curve with winding number ≥ 2 in euclidean plane $M = \mathbb{R}^2$. However, Theorem 11.5 holds for $n = 2$ (with the same proof) provided that the winding number of the curve is ± 1.

Remark 11.9 Theorem 11.4 is also false if the curvature can be negative. E.g. let S be the boundary of a small tubular neighborhood around a closed geodesic in a manifold M with $K < 0$. This is ε-convex (by the comparison theorem 3.1), but diffeomorphic to $S^1 \times S^{n-2}$, not to S^{n-1}.

Proof of Theorem 11.4. We are using the principle that in spaces with $K \geq 0$, we do not loose convexity by passing to interior parallel hypersurfaces (cf. Theorem 9.7).

Let S' be an embedded piece of $f(S)$. Then by 9.7, S' bounds (partially) some open set B on which $\rho = \rho_{S'}$ is concave. However, ρ is not strictly concave since it grows linearly along geodesics which are normal to S'. This can be improved similarly as in the proof of Toponogov's theorem (Ch.6): Instead of ρ consider

$$\sigma = \frac{1}{2}(R - \rho)^2. \tag{11.8}$$

where $R := 1/\varepsilon$. We compare with $\tilde{S}' = \partial B_R(0) \subset \mathbb{R}^n$; there we have

$$\tilde{\sigma}(x) = \frac{1}{2}(R - \tilde{\rho}(x))^2 = \frac{1}{2} < x, x >,$$

hence $D\nabla\tilde{\sigma} = I$. Therefore we get from the comparison theory (cf. 3.1)

$$D\nabla\sigma \geq I \tag{11.9}$$

in the sense of support functions (cf. Ch.9). Using convolution, we make σ smooth: Let

$$\sigma_\delta(x) = \int_{T_xM} \sigma(\exp_x(u))\varphi_\delta(\|u\|)du \tag{11.10}$$

where φ_δ is a mollifier with support in $B_\delta(0)$, i.e. $\varphi_\delta : \mathbb{R}_+ \to \mathbb{R}_+$ is smooth with $\varphi_\delta = const$ near 0, further $\varphi_\delta = 0$ on $[\delta, \infty)$ and $\int \varphi_\delta(\|x\|)d^nx = 1$. Now σ_δ is a smooth function with

$$|\sigma_\delta(x) - \sigma(x)| \leq \delta,\ \|\nabla\sigma_\delta\| \leq R,\ D\nabla\sigma_\delta \geq 1 \tag{11.11}$$

where an arbitrary small error (which goes to zero as $\delta \to 0$) is allowed in these estimates. (The second estimate comes from the Lipschitz constant of σ which is the maximum of $\frac{d}{d\rho}(\frac{1}{2}(R-\rho)^2)$ for $\rho > 0$.) Thus we have lost no convexity. In fact, we do not use σ_δ itself but a combination $\chi(\rho)\sigma + (1-\chi(\rho))\sigma_\delta$ for a suitable function χ; this agress to σ close to S' (for small ρ, smaller than the focal radius) and it agrees to σ_δ if ρ is big enough. This has the same good properties as σ_δ and additionally, it has S' as the level hypersurface with value $\frac{1}{2}R^2$. Let us keep the name σ_δ for this modified function.

If S were embedded, we could thus contract S (using the level sets of σ_δ) to a small ε-convex hypersurface lying in a small neighborhood of the origin of some exponential coordinate chart. Since $D \approx \partial$ near the origin for exponential coordinates (cf. Ch.1), this hypersurface is still ε-convex with respect to the euclidean geometry of the coordinate chart (for some smaller ε). So it bounds a disk (by 11.5).

However, σ and σ_δ are defined only in a small neighborhood of S'. But we may cover S with open subsets S_i such that $S_i' = f|S_i$ are embedded. Thus we recieve corresponding functions $(\sigma_\delta)_i$ defined near S_i' which we may past together along the immersed hypersurface $f(S)$: Note that the constant δ can be chosen independent of i, by compactness of S, and therefore $(\sigma_\delta)_i$ and $(\sigma_\delta)_j$ agree near $f(S_i \cap S_j)$. Now we may pass to the level hypersurfaces $\{(\sigma_\delta)_i = \frac{1}{2}(R-a)^2\}$ for some $a > 0$ independent of i which can be pasted together to an immersion $f_1 : S \to M$. This new immersion f_1 has roughly distance a from f and is still ε-convex; in fact it is ε_1-convex with $\varepsilon_1 = 1/R_1$ and $R_1 \approx R - a$. Moreover, the gradient lines of the local σ_δ's define a diffeomorphism between the hypersurfaces f and f_1 which is length decreasing by a factor $\leq e^{-a/R}$, due to the estimates (11.11). We will show that we can iterate the process with the *same* number a. Then, after N steps, $R_N \approx R - Na$ becomes arbitrarily small, so our immersion f_N gets arbitrarily small. So it is still ε-convex in some exponential coordinates and hence by Hadamard's theorem (Theorem 11.6), f_N is an embedding bounding a convex disk. Note that the whole process takes place in the relatively compact subset $M' = B_R(f(S))$.

However, in dimension $n = 2$, if we started with an ε-convex closed curve of winding number ≥ 2, we would get cusps and could not finish the contraction. The problem is that the embedded pieces S_i become smaller and smaller as we approach the cusp, and thus the values a must be chosen smaller and smaller. But we can exclude this behaviour in dimension $n \geq 3$.

Fig. 21.

Lemma 11.10 Let M^n, $n \geq 3$, be a complete manifold with $K \geq 0$ and $M' \subset M$ a relatively compact open subset. There exists a constant $r > 0$ depending only on ε and on the geometry of M' such that, for any compact $(n-1)$-dimensional manifold S, for any ε-convex immersion $f : S \to M'$, and for any $s \in S$, the connected component containing s of the set $f^{-1}(B_r(f(s))$ is embedded.

Proof. Let $f : S \to M'$ as above and fix $s \in S$. Using exponential coordinates around $f(s)$, we find some $\delta > 0$ (depending on ε and the geometry of M') such that $f|S'$ is still (say) $(\varepsilon/2)$-convex with respect to the euclidean metric on $B_{2\delta}(f(s))$ (in exponential coordinates), where S' denotes the connected component of s in the open subset $f^{-1}(B_{2\delta}(f(s)) \subset S$. Replacing $\varepsilon/2$ by ε again, we have to consider now an ε-convex immersion $f : S' \to \mathbb{R}^n$ with $f(s) = 0$ such that $f^{-1}(\overline{B_\delta(0)}) \subset S'$ is compact. We may assume that the unit normal vector at s is the n-th coordinate vector, $N(s) = e_n$. Now we intersect $f(S')$ by horizontal hyperplanes (parallel to the tangent plane at $f(s) = 0$), i.e. we look at the height function $h = x_n \circ f : S' \to \mathbb{R}$.

By $\varepsilon-$convexity, the only critical points of $x_n \circ f$ are local maxima or minima, and s is a local maximum with $h(s) = 0$. Consider the flow lines of $-\nabla h$ (with respect to the induced metric on S) starting at s. Either, they reach the boundary $\partial B_\delta(0)$, or they end at a local minimum. If no flow line reaches the boundary, they end all at the same minimum since $S' \setminus \{s\}$ is connected (here we need the dimension restriction $n - 1 \geq 2$).

Fig. 22.

In this case S' is compact (without boundary), so $S' = S$, and f is an embedding by Hadamard's theorem (cf. 11.5), and we are done.

So assume that some flow line of $-\nabla h$ reaches $\partial B_\delta(0)$. At which height is this possible? Along any of the flow lines, the heigth h decreases. So let $-r < 0$ be the highest h-value where some flow line reaches $\partial B_\delta(0)$. Then no flow line starting at s ends at a minimum of height $> -r$: otherwise, by connectedness of S', all flow lines would end at this minimum (they cannot reach the boundary before height $-r$) which we have excluded. Consider the connected component S'' of $\{h > -r\}$ which passes through s. Then still by Hadamard's theorem, $f|S''$ is an embedding, since the intersections of $f(S'')$ with the hyperplanes $\{x_n = t\} \cong \mathbb{R}^{n-1}$ for $0 > t > -r$ are closed ε-convex hypersurfaces in $\mathbb{R}^{n-1}$ (which have winding number ± 1 if $n-1 = 2$ since they contract to a point as $t \to 0$). To conclude the proof, we only have to find a good estimate for r.

Claim.

$$f(S'') \subset B_\delta(0) \cap \overline{B_R(-Re_n)} \tag{11.12}$$

where $R = 1/\varepsilon$, and consequently,

$$r \geq \bar{r} = \frac{1}{2}\delta^2\varepsilon. \tag{11.13}$$

Proof. To see (11.13) from (11.12), note that $f(S'')$ cannot reach $\partial B_\delta(0)$ before $\partial B_R(-Re_n)$. So let $-\bar{r}$ denote the height where $\partial B_R(-Re_n)$ meets $\partial B_\delta(0)$. Then

$$\delta^2 - \bar{r}^2 = R^2 - (R - \bar{r})^2 = 2\bar{r}R - \bar{r}^2,$$

hence $\bar{r} = \delta^2/(2R) = \frac{1}{2}\delta^2\varepsilon$.

Fig. 23.

Now we show (11.12). Put $S = f(S'') \subset \mathbb{R}^n$. (There is no danger of confusion with the previous S.) Note that $\bar{S}$ is a manifold with boundary $\partial S = \bar{S} \cap \{x_n = -r\}$. Let ρ be the signed distance function from S (continued negatively at the other side of S), defined on the open set $V \subset \mathbb{R}^n$ containing all points x where a shortest line segment from x to S exists, i.e. which are closer to S than to ∂S.

Fig. 24

As above, we put $\sigma = \frac{1}{2}(R - \rho)^2$. Then we have (cf. 11.9)

$$D\nabla\sigma \geq I$$

in the sense of support functions. (This holds also on the other side of S since we are now in euclidean space.) On the other hand, we consider also $\tilde{\sigma} = \frac{1}{2}(R - \tilde{\rho})^2$ where $\tilde{\rho}$ is the signed distance function from $\tilde{S} := \partial B_R(-Re_n)$. Hence $\tilde{\sigma}(x) = \|x - Re_n\|^2$ and

$$D\nabla\tilde{\sigma} = I.$$

Therefore $D\nabla(\sigma - \tilde{\sigma}) \geq 0$ which shows that $\sigma - \tilde{\sigma}$ is a convex function. Since $\sigma = \tilde{\sigma}$ and $\nabla\sigma = \nabla\tilde{\sigma}$ along $\mathbb{R}_+ \cdot e_n$, there are critical points (hence minima) for $\sigma - \tilde{\sigma}$ along $\mathbb{R}_+ \cdot e_n$. Thus we may conclude $(\sigma - \tilde{\sigma})(x) \geq 0$ for all $x \in V$ which can be connected to $\mathbb{R}_+ \cdot e_n$ by a straight line segment inside V. We must show that this holds in particular for all $x \in S$.

Let $S_1 \subset S$ be the set of points $x \in S$ where the vertical ray $x + \mathbb{R}_+ \cdot e_n$ meets S a second time and let $S_2 = S \setminus S_1$. All the points above S_1 are in V since they are closer to S_1 than to $\{x_n = -r\}$. Moreover, all points above S_2 are in V for the same reason.

Fig. 25.

Thus we may connect any point of S to some point in $\mathbb{R}_+ \cdot e_n$ within the shaded region (cf. fig. 26 below); we just have to avoid the cylinder of height r above ∂S if we start from S_2. This finishes the proof of the claim, of the lemma and of the theorem.

Fig. 26.

12. Lower Ricci curvature bounds and the Maximum Principle.

In this chapter, we want to discuss the following two theorems:

Theorem 12.1 (*Myers-Cheng* [31], [8])
Let M be complete with

$$Ric(M) \geq n - 1 = Ric(S^n).$$

Then $\operatorname{diam}(M) \leq \pi$, and equality holds if and only if M is isometric to S^n.

Theorem 12.2 (*Cheeger-Gromoll Splitting Theorem*, [7], [13])
Let M be complete with

$$Ric(M) \geq 0.$$

Then there exists a *line* in M, i.e. a complete geodesic which is shortest on any finite segment, if and only if M is isometric to a Riemannian product $M' \times \mathbb{R}$ for some complete $(n-1)$-manifold M' with $Ric(M') \geq 0$.

By definition, the *Riemannian product* $M = M_1 \times M_2$ of two Riemannian manifolds is the cartesian product with the metric

$$\|(v_1, v_2)\|^2 = \|v_1\|^2 + \|v_2\|^2$$

for any tangent vector (v_1, v_2) of $M_1 \times M_2$.

Let us first discuss the inequality of Theorem 12.1. This follows from our average comparison theorem 4.1: We saw that the first conjugate point on a geodesic γ on M, i.e. the first singularity t_1 for a Riccati solution A along γ with $A(t) \sim \frac{1}{t}I$ near $t = 0$, comes not later than on the comparison space S^n, i.e. at a distance $\leq \pi$. Beyond the first conjugate point, no geodesic can be shortest (cf. Ch.5), hence there are no shortest geodesics with length $> \pi$ and therefore the diameter is $\leq \pi$.

Corollary 12.3 A complete Riemannian manifold M with $Ric \geq n - 1$ is compact and has finite fundamental group.

Proof. Since $\operatorname{diam}(M) \leq \pi$, M is the image of the compact set $\overline{B_\pi(0)} \subset T_pM$ under the map $\exp_p$ (for any $p \in M$), hence it is compact. By the same reason, the universal cover $\hat{M}$ is compact. Since a covering map of compact spaces has only finitely many preimages (otherwise, the preimages would accumulate), the covering $\hat{M} \to M$ is finite, hence $\pi_1(M)$ is finite. ∎

Remark 12.4 Under the same hypothesis, we also have

$$\mathrm{Vol}(M) \leq \mathrm{Vol}(S^n).$$

with equality if and only if M is isometric to S^n. This follows immediately from the Bishop-Gromov inequality and its equality discussion.

Now let us come to the equality part of Theorem 12.1 and 12.2. These are *rigidity theorems*: the assumptions are so strong that we get a characterization up to isometries. The main ingredience is the Hopf-Calabi Maximum Principle for subharmonic functions:

Definition 12.5 Let M be any Riemannian manifold and $f : M \to \mathbb{R}$ a continuous function. Let $a \in \mathbb{R}$. We say

$$\Delta f \geq a$$

(*in the sense of support functions*) if for any $p \in M$ and any $\varepsilon > 0$ there is a smooth lower support function $\tilde{f} = f_{p,\varepsilon}$ (i.e. $\tilde{f} \leq f$, $\tilde{f}(p) = f(p)$) defined in a neighborhood of p with

$$\Delta \tilde{f}(p) \geq a - \varepsilon$$

where $\Delta \tilde{f} = \mathrm{trace} D\nabla \tilde{f} = \mathrm{div}\nabla \tilde{f}$ is the Riemannian *Laplace operator* (*Laplace- Beltrami operator*. Similarly we define $\Delta f \leq a$ using smooth upper support functions. Clearly, $\Delta f \leq a$ iff $\Delta(-f) \geq -a$. A continuous function f with $\Delta f \geq 0$ is called *subharmonic*, and if $\Delta f \leq 0$, f is called *superharmonic*.

Theorem 12.6 (*Hopf-Calabi Maximum Principle*, [23], [4], [13])
Let M be a connected Riemannian manifold and $f : M \to \mathbb{R}$ a continuous subharmonic function. Then f attains no maximum unless it is constant.

Proof. If f attains a maximum at $p \in M$ and is not constant on any neighborhood of p, we may choose a small coordinate ball U around p such that

$$\partial' U := \{x \in \partial U;\ f(x) = f(p)\}$$

is a proper subset of ∂U. Now pick a smooth function with
(a) $h(p) = 0$,
(b) $h < 0$ on U,
(c) $\Delta h \geq 0$ on U.
In fact, h can be constructed easily in the form

$$h = e^{\alpha\phi} - 1$$

for some function ϕ and a sufficiently large constant $\alpha > 0$ since

$$\Delta(e^{\alpha\phi} - 1) = (\alpha^2 \|\nabla\phi\|^2 + \alpha\Delta\phi)e^{\alpha\phi}.$$

If $\eta > 0$ is sufficiently small, we have $f + \eta h < f(p)$ on ∂U while $(f + \eta h)(p) = f(p)$. This shows that $f + \eta h$ attains a maximum on U, say at q. Then also the lower support function $f_{q,\varepsilon} + \eta h$ of $f + \eta h$ at q takes a maximum at q, but

$$\Delta(f_{q,\varepsilon} + \eta h)(q) \geq -\varepsilon + \eta\Delta h(q) > 0$$

if we choose ε sufficiently small. This is a contradiction since the Hessian of a function at a maximum point is negative semidefinite, so its trace is ≤ 0.
Thus, the set of points where f attains a maximum is open and closed and by hypothesis not empty, hence it is the whole manifold. ■

Proof of Theorem 12.1, equality part:
Let $p, q \in M$ with $|p,q| = \pi$. Let $\rho_p(x) = |x,p|$, $\rho_q(x) = |x,q|$, and $f = \rho_p + \rho_q - \pi$. By triangle inequality we have $f \geq 0$, and $f = 0$ on any shortest geodesic γ from p to q.

Fig. 27.

Moreover, due to $Ric \geq n-1$ and the average comparison theorem 4.1, we have on $M \setminus \{p,q\}$:

$$\Delta\rho_p \leq (n-1)\cot\rho_p, \quad \Delta\rho_q \leq (n-1)\cot\rho_q$$

in the sense of support functions. In fact, to prove the first inequality at some point $x \in M \setminus \{p,q\}$, we choose a shortest geodesic segment β from x to p and replace p by some point p' on β close to p; then the distance function $\rho_{p'}$ from p' is smooth near x and satisfies the above inequality with an arbitrary small error (by Theorem 4.1), and by trangle inequality, $\rho_{p'} + |p', p|$ is an upper support function for ρ_p. (A similar argument was used in the proof of Toponogov's Theorem 6.1, Case 2.)

Since $\rho_q \geq \pi - \rho_q$, we have

$$\cot \rho_q \leq \cot(\pi - \rho_p) = -\cot \rho_p.$$

Thus

$$\Delta f \leq (n-1)(\cot \rho_p + \cot \rho_q) \leq 0$$

(in the sense of support functions) on $M \setminus \{p, q\}$. Since f attains the maximum 0 on γ, the maximum principle applied to $-f$ gives $f \equiv 0$. Consequently, $\rho_q = \pi - \rho_p$, and any geodesic starting from p meets q at the distance π. Now the equality discussion of the average comparison theorem 4.1 shows $R_V = I$ where V is the radial vector field from p, and hence M is isometric to S^n (where the isometry is via $\exp_p$). ∎

Fig. 28.

The proof of Theorem 3 is quite similar. We first show the superharmonicity of the Busemann functions which corresponds to the concavity in the case of $K \geq 0$ (cf. Lemma 10.6):

Lemma 12.7 Let M be complete with $Ric \geq 0$ and $\gamma : [0, \infty) \to M$ be any ray in M. Then the corresponding Busemann function b_γ is superharmonic, i.e. $\Delta b_\gamma \leq 0$ in the sense of support functions.

Proof. Recall that for any $x \in M$, we have a smooth upper support function

$$b_{x,t}(y) := |\gamma_x(t), y| - t + b_\gamma(x)$$

of b_γ at x, where γ_x is an asymptotic ray starting at x (cf. Ch.10). By the avarage comparison theorem 4.1 we have

$$\Delta b_{x,t}(x) \le \frac{n-1}{t},$$

thus $\Delta b_\gamma \le 0$ in the sense of support functions. ■

Proof of Theorem 12.2. (cf. [13])

Devide the line γ in M into the two rays $\gamma^+, \gamma^- : [0,\infty) \to M$ by putting $\gamma^\pm(t) = \gamma(\pm t)$, $t \ge 0$. Let $b^\pm$ be the Busemann function associated to $\gamma^\pm$. Then $b^+ + b^- \ge 0$ by triangle inequality, with equality along γ. Moreover, by the previous lemma we have

$$\Delta(b^+ + b^-) \le 0$$

in the sense of support functions. Thus $b^+ + b^- \equiv 0$ by the maximum principle. Moreover, for any $x \in M$,

$$b^+_{x,t} \ge b^+ = -b^- \ge -b^-_{x,t}. \tag{12.1}$$

Thus, $b := b^+$ is once differentiable at x, and the asymptotic rays γ_x^+, γ_x^- fit together to a complete geodesic γ_x perpendicular to the level hypersurface $\{y \in M;\ b(y) = b(x)\}$. Further, let

$$A_t^\pm(u) = D\nabla b^\pm_{x,t}(\gamma_x(u)).$$

For fixed u, $A_t^\pm(u)$ is monotonely decreasing with t, and bounded below since $A_t^+(u) \ge -A_t^-(u)$ by (12.1). Thus $A_t^\pm$ converges to some solution $A^\pm$ of the Riccati equation along γ_x. Moreover, $A^+ \le -A^-$ and trace$A^\pm \le 0$ which implies $A^+ = -A^- =: A$ with trace$A = 0$. The Riccati equation gives

$$\|A\|^2 = \mathrm{trace} A^2 = -Ric(\gamma_x') \le 0,$$

thus $A \equiv 0$. Hence $D\nabla b^\pm_{x,t}(x) \to 0$ as $t \to \infty$, and therefore, $b^\pm$ is concave (cf. Ch.10). Since $b^+ = -b^-$, b is also convex and hence *affine*, i.e. $b \circ \beta = 0$ for any geodesic β. Thus b is smooth with $D\nabla b = 0$, i.e. ∇b is a *parallel vector field.* Let M' denote the level hypersurface $\{b = 0\}$ and ϕ_t be the flow of ∇b. Then the map $\Phi : M' \times \mathbb{R} \to M$ with

$$\Phi(x,t) = \phi_t(x) = \gamma_x(t)$$

is an isometry. ■

There are interesting applications of Theorem 12.2 since sometimes, we get a line for free:

Corollary 12.8 Let M be a complete non-compact *irreducible* manifold (i.e. not a Riemannian product) such that $Ric \geq 0$. Then M has only one end, for any compact subset $C \subset M$, the complement $M \setminus C$ has only one unbounded connected component.

(By definition, an *end* of a non-compact manifold M is a function $E : \mathcal{K} \to \mathcal{M}$, where $\mathcal{K}$ is the set of all compact subsets of M and $\mathcal{M}$ the open ones, such that $E(C)$ is a connected component of $M \setminus C$ with $E(D) \subset E(C)$ whenever $D \supset C$.)

Proof. If we had two different unbounded connected components of $M \setminus C$, we would take diverging sequences (p_i), (q_i) in each component and join p_i to q_i by a shortest geodesic segment γ_i. These geodesics have to pass through C, so they accumulate, and a limit geodesic is a line (since $|p_i, q_i| \to \infty$) which is excluded by irreducibility and Theorem 12.2. ■

Corollary 12.9 Let M be a compact manifold with $Ric \geq 0$. Let us suppose that the universal cover $\hat{M}$ is irreducible. Then $\pi_1(M)$ is finite.

Proof. Assume $\hat{M}$ non-compact (otherwise we are done). Then there is a compact fundamental domain F related to $\Gamma \cong \pi_1(M)$ (acting isometrically on $\hat{M}$). Fix $o \in F$. Choose a sequence $p_i \to \infty$ in M and join o to p_i by a shortest geodesic segment γ_i. Let q_i be the midpoints of γ_i. Since F is a fundamental domain, there exist $g_i \in \Gamma$ such that $g_i q_i \in F$. Thus, as in the proof of Corollary 12.8, the shortest geodesic segments $g_i \gamma_i$ accumulate to a line γ. This is impossible by Theorem 12.2 and the irreducibility of $\hat{M}$. ■

13. The Bochner technique.

Let M be a Riemannian manifold and consider a vector field V on M. In Ch.2, we considered the derivative $A = DV$ and derived the equation

$$D_V A + A^2 + R_V = DW$$

with $W = D_V V$ (cf. (2.5)). Taking the trace, we get

$$\partial_V \text{div}(V) + \text{trace}(A^2) + Ric(V) = \text{div}(W), \tag{13.1}$$

(Recall that the *divergence* of a vector field X is given by $\text{div}(X) = \text{trace}DX$.) Suppose now M is compact with no boundary, then

$$\int_M (\partial_V \text{div}(V) + \text{trace}(A^2) + Ric(V)) = 0. \tag{13.2}$$

by the Divergence Theorem (since $\partial M = \emptyset$).

Theorem 13.1 (*Bochner*)
Let M be a compact manifold without boundary.

a) If A is symmetric, $\text{div}(V) = 0$ and $Ric(V) \geq 0$, then

$$DV = 0 \quad \text{and } Ric(V) = 0.$$

b) If A is antisymmetric (in particular, $\text{div}(V) = 0$) and $Ric(V) \leq 0$, then

$$DV = 0 \quad \text{and } Ric(V) = 0.$$

Proof. This is clear by (13.2): Since the first term vanishes by hypothesis, the integrand does not change sign, so it must vanish pointwise.

Corollary 13.2 Let M be compact with $Ric(V) > 0$. Then the first Betti number $b_1(M)$ vanishes.

Proof. Since the first de-Rham cohmomology of M is

$$H^1(M) = \{\text{local gradients}\}/\{\text{gradients}\},$$

we only have to show that each *local gradient vector field* V (i.e. $DV = (DV)^*$, cf. Ch.2) is the gradient of some function. Put $\tilde{V} = V - \nabla f$ where $f : M \to \mathbb{R}$ is a solution of the equation $\Delta f = \text{div}(V)$. Then $\tilde{V}$ is still a local gradient, and moreover, $\text{div}(\tilde{V}) = 0$. Thus Theorem 13.1 a) applies to $\tilde{V}$, and now we get the contradiction $Ric(\tilde{V}) = 0$, unless $\tilde{V} = 0$, i.e. $V = \nabla f$, so V is a gradient. ∎

Corollary 13.3 Let M be compact with $Ric < 0$. Then the group of isometries of M is finite.

Proof. It is known (cf. [26]) that $I(M)$, the *group of isometries* of M, is a compact Lie group. Any one-parameter subgroup $(g_t)_{t\in\mathbb{R}}$ of $I(M)$ gives rise to a vector field V on M, defined by

$$V(x) = \frac{d}{dt} g_t(x)|_{t=0}$$

which is the *Killing field* corresponding to (g_t). Since g_t is an isometry,

$$\langle (dg_t)_x.a, (dg_t)_x.b\rangle = \langle a, b\rangle$$

for all $a, b \in T_x M$. Differentiating with respect to t we see that $\langle D_a V, b\rangle + \langle a, D_b V\rangle = 0$, i.e. $A = DV$ is skew symmetric. (In fact, this property characterizes Killing fields.) Now from Theorem 13.1 b) we get $Ric(V) = 0$ which is a contradiction to $Ric < 0$ unless $V = 0$. So there are no nontrivial one parameter subgroups in $I(M)$. Hence $I(M)$ is discrete; since it is also compact, it must be finite. ∎

Remark 13.4 Corollaries 13.2 and 13.3 can be easily extended to the case where $Ric \geq 0$ or $Ric \leq 0$. Then applying 13.1 to a divergence-free local gradient field (Case (a)) or to a Killing field (Case (b)) V, we recieve $DV = 0$. In Case (a) we recieve that the first Betti number is the number of linear independent parallel vector fields on M, so in particular, $b_1(M) \leq n$, and in Case (b) we get that any Killing field is parallel, so the connected component of $I(M)$ acts only on a flat factor of M.

Remark 13.5 It is interesting to compare the proofs of Corollaries 13.1 (Bochner) and 12.2 (Myers) which both show that a compact manifold M with $Ric > 0$ has $b_1(M) = 0$. (Recall that by Hurewitz, $b_1(M) = 0$ if $\pi_1(M)$ is finite, but the converse does not hold.) In both proofs we have used the equation

$$\partial_V \mathrm{trace}(A) + \mathrm{trace}(A^2) + Ric(V) = \mathrm{div}(W)$$

for a local gradient vector field V, where $A = DV$ and $W = D_V V$. However, in Myers' theorem, we have assumed $W = 0$ and estimated trace$A$ (as a solution of the Riccati equation) while in Bochner's theorem, we have assumed trace$A = 0$ and allowed arbitrary W. Myers' result is stronger for $Ric > 0$, however, Bochner's technique gives a result also for $Ric \geq 0$.

Appendix: Nested Coverings

In this appendix, we want to prove the topological result on coverings which we have used in Ch.8, cf. (8.20). The following exposition is essentially due to U. Abresch (cf. [28]). Let us first recall the *Mayer-Vietoris principle* (cf [3]). Let $X = X_1 \cup ... \cup X_N$ be a topological space. Let SX denote the complex of singular chains $c = \sum_i \alpha_i \sigma_i$ where $\sigma_i : \Delta^q \to X$ are singular (q-)simplices and α_i coefficients in the chosen field $\mathbf{F}$. The homology of X can be computed from the subcomplex $\tilde{S}X$ of *small* chains where a chain c is called small if all the σ_i take values in some X_j. For $N = 2$ we get the usual Mayer-Vietoris exact sequence of chain complexes:

$$0 \to S(X_1 \cap X_2) \to SX_1 \oplus SX_2 \to \tilde{S}X \to 0$$
$$(c_1, c_2) \mapsto c_1 + c_2$$
$$c \mapsto (c, -c)$$

For arbitrary N, we put
$$X_{i_1,...,i_k} := X_{i_1} \cap ... \cap X_{i_k},$$
and putting $C_0 = \tilde{S}X$, $C_1 = \sum_i SX_i$, $C_2 = \sum_{i,j} X_{ij}$ etc., we get an exact sequence

$$\to C_3 \to C_2 \to C_1 \to C_0 \to 0$$

where the maps $\delta_k : C_k \to C_{k-1}$ are defined by their "matrix elements"

$$\delta_k(i_1, ..., i_k; j_1, ..., j_{k-1}) : SX_{i_1,...,i_k} \to SX_{j_1,...,j_{k-1}}$$

as follows: the only nonzero matrix elements are

$$\delta_k(i_1, ..., i_k; i_1, ...\hat{i_j}..., i_k) = (-1)^{j-1} \cdot inc$$

where *inc* denotes the natural inclusion map. Put

$$A_k = \mathrm{im}\delta_k = \ker\delta_{k-1} \subset C_{k-1}.$$

In particular, we have
$$A_1 = C_0 = \tilde{S}X.$$

The above exact sequence of chain complexes can be split into short exact sequences as follows:
$$0 \longrightarrow A_{k+1} \longrightarrow C_k \xrightarrow{\delta_k} A_k \longrightarrow 0.$$

As usual, a short exact sequence of chain complexes gives rise to a long exact sequence of the homologies:

$$\to H_pA_{k+1} \to H_pC_k \to H_pA_k \to H_{p-1}A_{k+1} \to ...$$

We will use only the segment

$$H_pC_k \to H_pA_k \to H_{p-1}A_{k+1}$$

in order to estimate the middle term. If we had only to compute the Betti numbers, we would get

$$\text{dim}H_pA_k \leq \text{dim}H_pC_k + \text{dim}H_{p-1}A_{k+1} \qquad (A1)$$

for all k, and in particular

$$b_p(X) = \text{dim}H_pA_1 \leq \text{dim}H_pC_1 + \text{dim}H_{p-1}A_2,$$

and further

$$\text{dim}H_{p-1}A_2 \leq \text{dim}H_{p-1}C_2 + \text{dim}H_{p-2}A_3$$

and so on, hence by induction

$$b_p(X) \leq \text{dim}H_pC_1 + \text{dim}H_{p-1}C_1 + ... + \text{dim}H_0C_{p+1}$$

since $H_{p-q}A_{q+1} = 0$ for $q > p$. However, these Betti numbers are not available in our application; instead, we have to compute the *content* which is the *rank* of certain inclusion maps. Unfortunately, the analogue of (A1) for the rank in place of the dimension is not true: If we have a commutative diagram of exact sequences,

$$\begin{array}{ccccc} A & \xrightarrow{\phi} & B & \xrightarrow{\psi} & C \\ \downarrow \alpha & & \downarrow \beta & & \downarrow \gamma \\ A' & \xrightarrow{\phi'} & B' & \xrightarrow{\psi'} & C' \end{array}$$

we have *not* $\text{rk}(\beta) \leq \text{rk}(\alpha) + \text{rk}(\gamma)$; e.g. we could choose $\alpha = 0$ and $\gamma = 0$, and β maps $\text{im}(\phi)$ onto 0 while a complement of $\text{im}\phi$ is mapped into $\text{ker}(\psi')$. However, if we have three such sequences,

$$\begin{array}{ccccc} A & \xrightarrow{\phi} & B & \xrightarrow{\psi} & C \\ \downarrow \alpha & & \downarrow \beta & & \downarrow \gamma \\ A' & \xrightarrow{\phi'} & B' & \xrightarrow{\psi'} & C' \\ \downarrow \alpha' & & \downarrow \beta' & & \downarrow \gamma' \\ A'' & \xrightarrow{\phi''} & B'' & \xrightarrow{\psi''} & C'' \end{array}$$

then we get

$$\mathrm{rk}(\beta' \circ \beta) \leq \mathrm{rk}(\alpha') + \mathrm{rk}(\gamma).$$

Proof. Consider a decomposition

$$\mathrm{im}(\beta) = (\mathrm{im}\beta)_1 \oplus (\mathrm{im}\beta)_2$$

where

$$(\mathrm{im}\beta)_1 = \mathrm{im}(\beta) \cap \ker(\psi').$$

Then $(\mathrm{im}\beta)_1 \subset \mathrm{im}(\phi')$, hence

$$\beta'(\mathrm{im}\beta)_1 \subset \phi''(\alpha'(A')).$$

Moreover, ψ' is injective on $(\mathrm{im}\beta)_2$, and

$$\psi'(\mathrm{im}\beta)_2 \subset \gamma(\psi(B)).$$

Therefore,

$$\begin{aligned} \mathrm{rk}(\beta' \circ \beta) &\leq \dim(\beta'(\mathrm{im}\beta)_1) + \dim(\beta'(\mathrm{im}\beta)_2) \\ &\leq \quad \mathrm{rk}(\alpha') \qquad + \quad \mathrm{rk}(\gamma) \end{aligned}$$

■

We apply this to

$$\begin{array}{ccccc} H_pC_k^0 & \longrightarrow & H_pA_k^0 & \longrightarrow & H_{p-1}A_{k+1}^0 \\ \downarrow & & \downarrow & & \downarrow \\ H_pC_k^p & \longrightarrow & H_pA_k^p & \longrightarrow & H_{p-1}A_{k+1}^p \\ \downarrow & & \downarrow & & \downarrow \\ H_pC_k^{p+1} & \longrightarrow & H_pA_k^{p+1} & \longrightarrow & H_{p-1}A_{k+1}^{p+1} \end{array}$$

where C_k^p and A_k^p are the above defined complexes, but now for p+2 different coverings $\{X_1^q, ..., X_N^q\}$ of X $(q = 1, ..., p+2)$ with the property that

$$X_i^{q-1} \subset X_i^q$$

for $q = 1, ..., p+1$ (*nested coverings*). Thus we get

$$\begin{aligned} &\mathrm{rk}(H_pA_k^0 \to H_pA_k^{p+1}) \\ &\qquad \leq \mathrm{rk}(H_pC_k^p \to H_pC_k^{p+1}) + \mathrm{rk}(H_{p-1}A_{k+1}^0 \to H_{p-1}A_{k+1}^p) \end{aligned}$$

for all k and in particular

$$\begin{aligned}&\operatorname{rk}(H_pA_1^0 \to H_pA_1^{p+1})\\&\quad\leq \operatorname{rk}(H_pC_1^p \to H_pC_1^{p+1}) + \operatorname{rk}(H_{p-1}A_2^0 \to H_{p-1}A_2^p)\end{aligned}$$

and further

$$\begin{aligned}&\operatorname{rk}(H_{p-1}A_2^0 \to H_{p-1}A_2^{p+1})\\&\quad\leq \operatorname{rk}(H_{p-1}C_2^{p-1} \to H_{p-1}C_2^p) + \operatorname{rk}(H_{p-2}A_3^0 \to H_{p-2}A_3^{p-1})\end{aligned}$$

and so on. Thus we get by induction

$$\begin{aligned}&\operatorname{rk}(H_pA_1^0 \to H_pA_1^{p+1})\\&\quad\leq \sum_{j=0}^{p} \operatorname{rk}(H_{p-j}C_{j+1}^{p-j} \to H_{p-j}C_{j+1}^{p+1-j})\end{aligned} \qquad (A2)$$

Recall that

$$\begin{aligned}&H_qC_K^q \to H_qC_k^{q+1}\\&\quad= \bigoplus_{i_1<\ldots<i_k} (H_q(X_{i_1}^q \cap \ldots \cap X_{i_k}^q) \to H_q(X_{i_1}^{q+1} \cap \ldots \cap X_{i_k}^{q+1}).\end{aligned}$$

We apply this to our coverings of balls $B_i^j = 10^j B_i$ for $j = 0, ..., n+1$ and recieve

$$\begin{aligned}&\operatorname{cont}(\bigcup_i B_i, \bigcup_i 10^{n+1}B_i)\\&\quad= \sum_{p=0}^{n} \operatorname{rk}(H_p(\bigcup_i B_i^0) \to H_p(\bigcup_i B_i^{n+1}))\\&\quad\leq \sum_{p=0}^{n} \operatorname{rk}(H_p(\bigcup_i B_i^{n-p}) \to H_p(\bigcup_i B_i^{n+1}))\end{aligned}$$

We estimate each term individually by (A2): For given p we put

$$X = \bigcup_i B_i^{n-p}\ , \quad X_i^q = B_i^{n-p+q}$$

and sum over p. Putting now

$$C_k^q = \sum_{i_1<\ldots<i_k} S(B_{i_1}^q \cap \ldots \cap B_{i_k}^q)$$

and

$$\gamma^q_{p,k} : H_p C^q_k \to H_p C^{q+1}_k,$$

we get from (A2)

$$\begin{array}{lllllllllll}
\multicolumn{11}{l}{\text{cont}(\bigcup_i B_i, \bigcup_i 10^{n+1} B_i)} \\
\leq & & \text{rk } \gamma^n_{n,1} & + & \text{rk } \gamma^{n-1}_{n-1,2} & + & ... & + & \text{rk } \gamma^1_{1,n} & + & \text{rk } \gamma^0_{0,n+1} \\
 & + & \text{rk } \gamma^n_{n-1,1} & + & \text{rk } \gamma^{n-1}_{n-2,2} & + & ... & + & \text{rk } \gamma^1_{0,n} & & \\
 & + & ... & & & & & & & & \\
 & + & \text{rk } \gamma^n_{1,1} & + & \text{rk } \gamma^{n-1}_{0,2} & & & & & & \\
 & + & \text{rk } \gamma^n_{0,1} & & & & & & & & \\
\leq & & \text{rk } \gamma^n_{*,1} & + & \text{rk } \gamma^{n-1}_{*,2} & + & ... & + & \text{rk } \gamma^1_{*,n} & + & \text{rk } \gamma^0_{*,n+1}
\end{array}$$

Hence

$$\begin{aligned}
&\text{cont}(\bigcup_i B_i, \bigcup_i 10^{n+1} B_i) \\
&\leq \sum_{k=1}^{n+1} \text{rk } \gamma^{n+1-k}_{*,k} \\
&= \sum_{k=1}^{n+1} \sum_{i_1<\ldots<i_k} \text{cont}(\bigcap_{j=1}^{k} B^{n+1-k}_{i_j}, \bigcap_{j=1}^{k} B^{n+2-k}_{i_j})
\end{aligned}$$

which proves Equation (8.20).

References:

[1] Abresch, U. *Lower curvature bounds, Toponogov's theorem and bounded topology I,II.* Ann. scient. Éc. Norm. Sup., 4^e série, t.18 (1985), 651-670

[2] Berger, M.: *Les variétés riemanniennes (1/4)-pincées.* Ann. Sc. Norm. Super. Pisa, III 14 (1960), 161-170

[3] Bott, R., Tu, L.W.: *Differential forms in algebraic topology.* Springer Graduate Texts in Mathematics 82 (1982)

[4] Calabi, E.: *An extension of E.Hopf's maximum principle with an application to Riemannian geometry.* Duke Math. J. 25 (1957), 45-56

[5] Cheeger, J., Ebin, D.G.: *Comparison theorems in Riemannian geometry* American Elsevier, New York 1975

[6] Cheeger, J., Gromoll, D.: *On the structure of complete manifolds of nonnegative curvature.* Ann. of Math. 96 (1972), 413-442

[7] Cheeger, J., Gromoll, D.: *The splitting theorem for manifolds of nonnegative Ricci curvature.* J. Diff. Geom. 6 (1971), 119-128

[8] Cheng, S.Y.: *Eigenvalue comparison theorem and its geometric application.* Math. Z. 143 (1975), 289-297

[9] Eschenburg, J.-H. *Comparison theorems and hypersurfaces.* Manuscripta math., 59 (1987), 295-323

[10] Eschenburg, J.-H. *Diameter, volume, and topology for positive Ricci curvature.* Journal of Differential Geometry, 33 (1991), 743-747

[11] Eschenburg, J.-H. *Local convexity and nonnegative curvature - Gromov's proof of the sphere theorem.* Invent. math., 84 (1986), 507-522

[12] Eschenburg, J.-H. *Maximum Principle for Hypersurfaces*, Manuscripta Math. 64 (1989), 55 - 75

[13] Eschenburg, J.-H., Heintze, E. *An elementary proof of the Cheeger - Gromoll splitting theorem.* Ann. Glob. Analysis and Geometry, vol. 2, n. 2 (1984), 141-151

[14] Eschenburg, J.-H., Heintze, E. *Comparison theory for Riccati equation* Manuscripta math., 68 (1990), 209-214

[15] Green, L.W.: *A theorem of E.Hopf.* Mich. Math. J. 5 (1958), 31-34

[16] Greene, R.E., Wu, H.: *On the subharmonicity and plurisubharmonicity of geodesically convex functions.* Ind. Univ. Math. J. 22 (1973), 641-653

[17] Greene, R.E., Wu, H.: *C^∞-convex functions and manifolds of positive curvature.* Acta Math. 137 (1976), 209-245

[18] Gromoll, D., Klingenberg, W., Meyer, W.: *Riemannsche Geometrie im Großen* Lecture notes 55, Springer 1968

[19] Gromov, M.: *Curvature, diameter and Betti numbers.* Comm. Math. Helv. 56 (1981), 179-195

[20] Gromov, M.: *Structures métriques pour les variétés riemanniennes.* Cedic-Nathan 1981

[21] Grove, K., Shiohama, A.: *A generalized sphere theorem.* Ann. of Math. 106 (1977), 201-211

[22] Hadamard, J.: *Sur certain propriétés des trajectories en dynamique.* J. Math. Pures Appl. (5)3 (1897), 331-387

[23] Hopf, E.: *Elementare Bemerkungen über die Lösungen partieller Differentialgleichungen 2.Ordnung vom elliptischen Typ.* Sitzungsber. Preuss. Ak. d. Wiss. 19 (1927), 147-152

[24] Karcher, H.: *Riemannian comparison constructions.* In: S.S.Chern (ed.): Studies in Global Geometry and Analysis, M.A.A. Studies in Mathematics, vol. 27 (1989)

[25] Klingenberg, W.: *Über Riemannsche Mannigfaltigkeiten mit positiver Krümmung.* Comm. Math. Helv. 35 (1961), 47-54

[26] Kobayashi, S., Nomizu, K.: *Foundations of differential geometry, I.* Interscience, Wiley, New York 1963

[27] Kürzel, M. *Der Vergleichssatz von Toponogov im Fall variabler Krümmung.* Diplomarbeit, Münster 1994

[28] Meyer, W. *Toponogov's theorem and applications.* Preprint University of Münster (1990)

[29] Milnor, J.: *Morse Theory.* Princeton 1962

[30] Milnor, J.: *A note on curvature and fundamental group.* J. Diff. Geom. 2 (1968), 1-7

[31] Myers, S.B.: *Riemannian manifolds with positive mean curvature.* Duke Math. J. 8 (1941), 401-404

[32] Sharafutdinov, V.A.: *Convex sets in a manifold of nonnegative curvature.* Math.

Zametki 26 (1979) 556-560

[33] Strake, M., Walshap, G. : Σ-*flat manifolds and Riemannian submersions.* Manuscr. Math. 64 (1989), 213-226

[34] Wu, H.: *An elementary method in the study of nonnegative curvature.* Acta Math. 142 (1979), 57-78

[35] J.-W.Yim: *Distance nonincreasing retraction on a complete open manifold of nonnegative sectional curvature.* Ann. Glob. Anal. and Geom.

Index

J.-H. Eschenburg:

Comparison Theorems in Riemannian Geometry

Figures

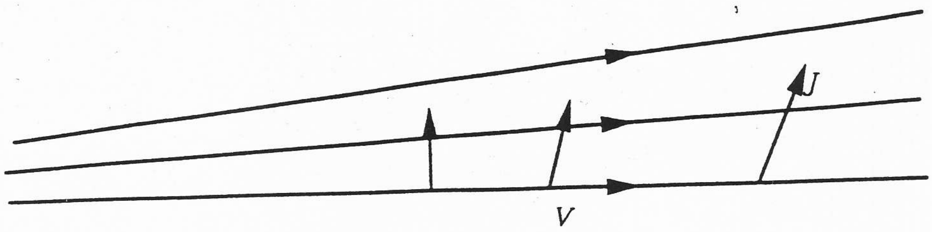

Fig. 1.

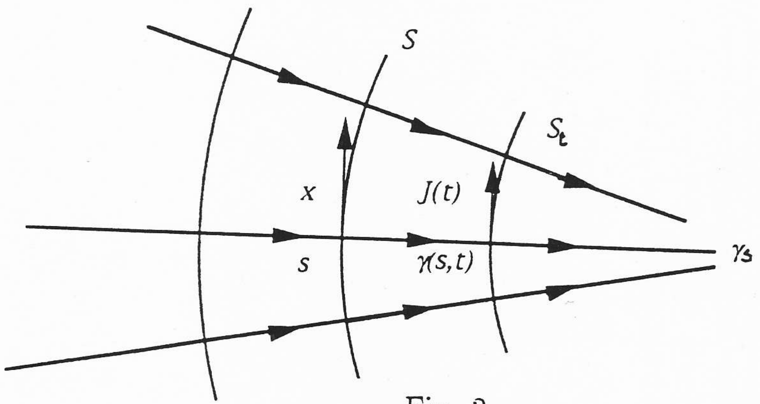

Fig. 2.

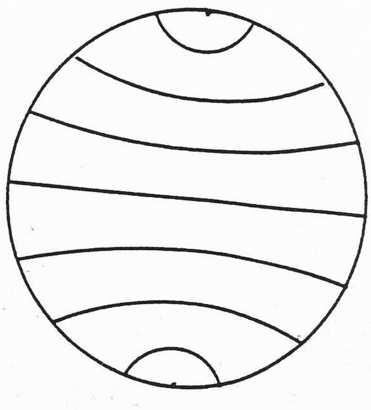

Fig. 3.

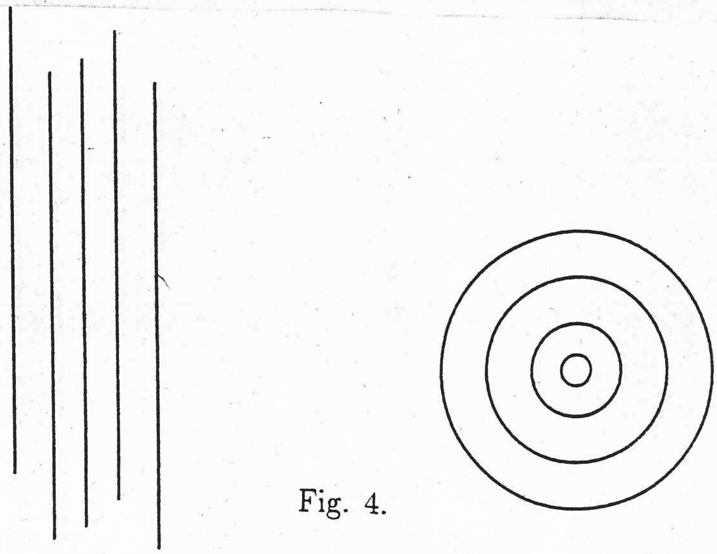

Fig. 4.

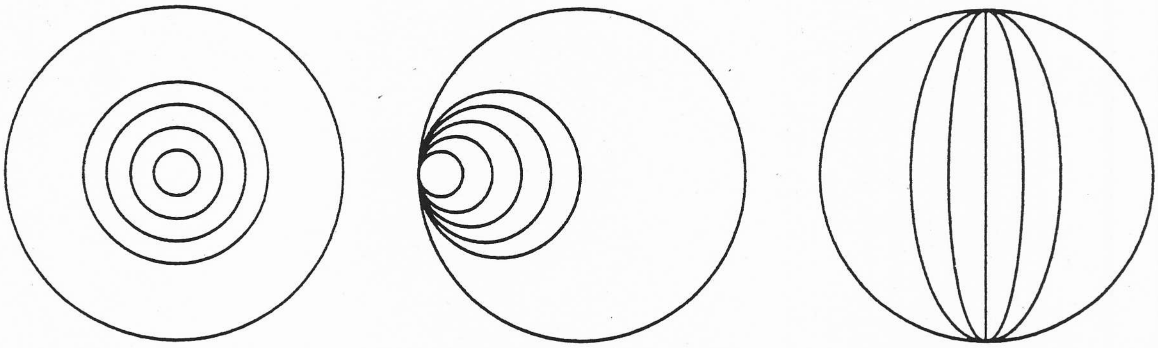

Fig. 5.

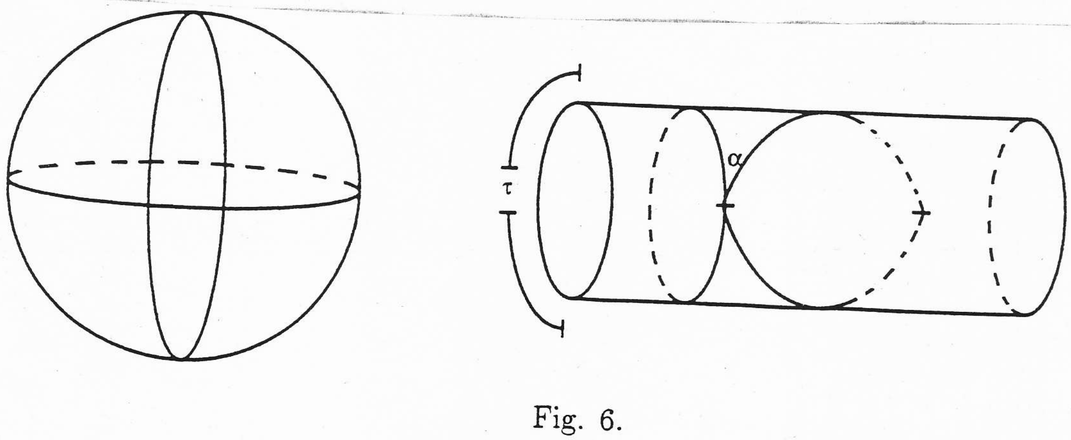

Fig. 6.

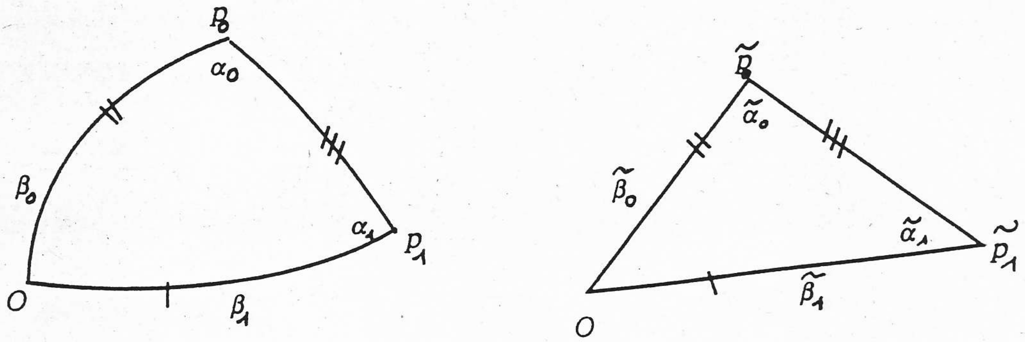

Fig. 7.

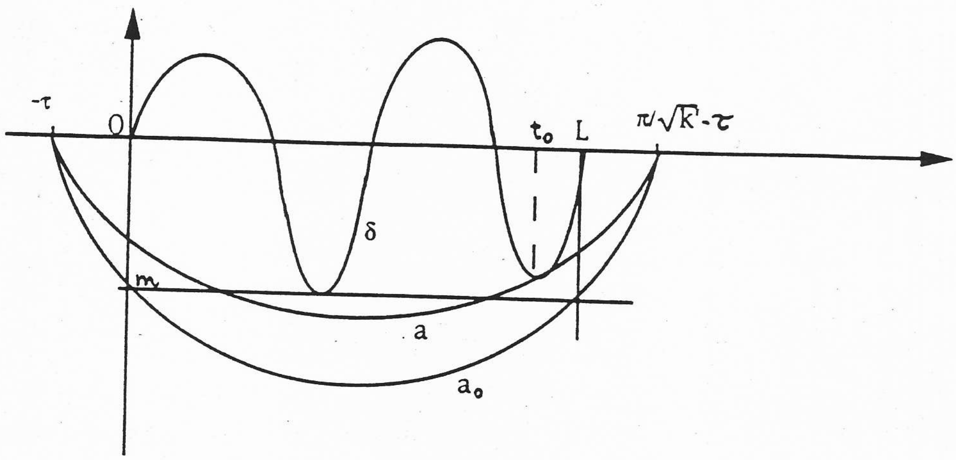

Fig. 8.

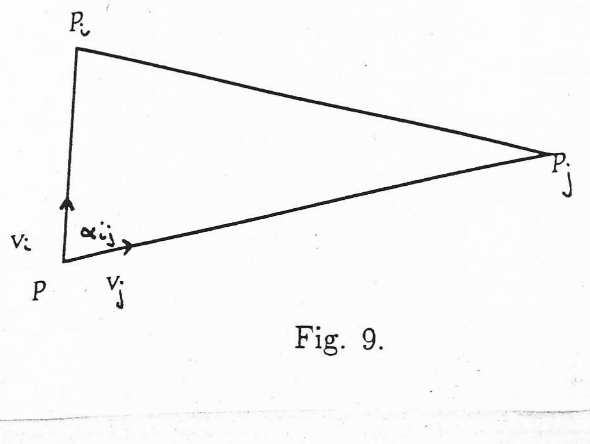

Fig. 9.

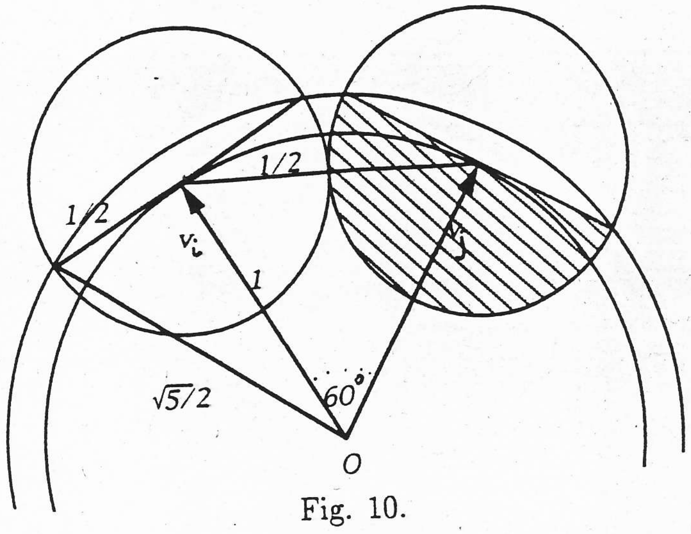

Fig. 10.

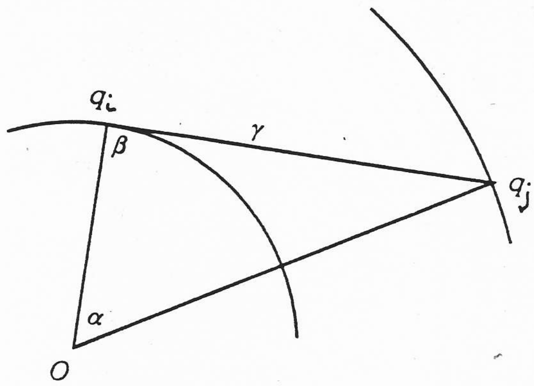

Fig. 11.

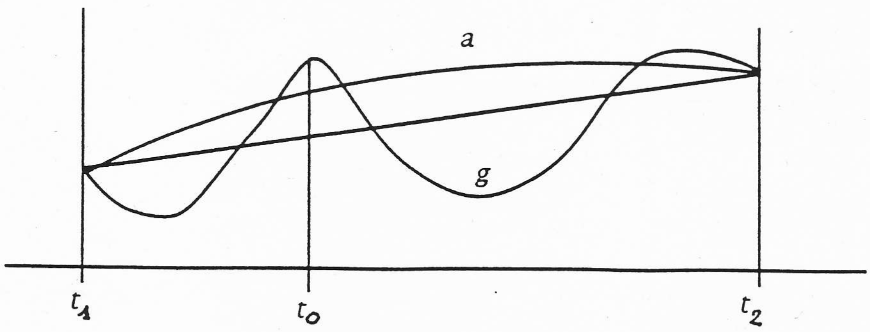

Fig. 12.

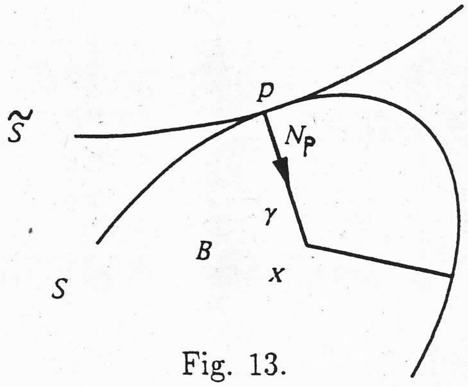

Fig. 13.

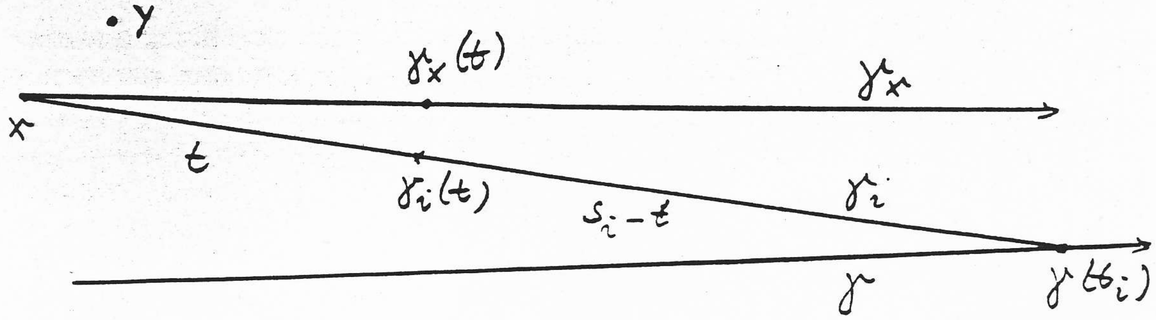

Fig. 14.

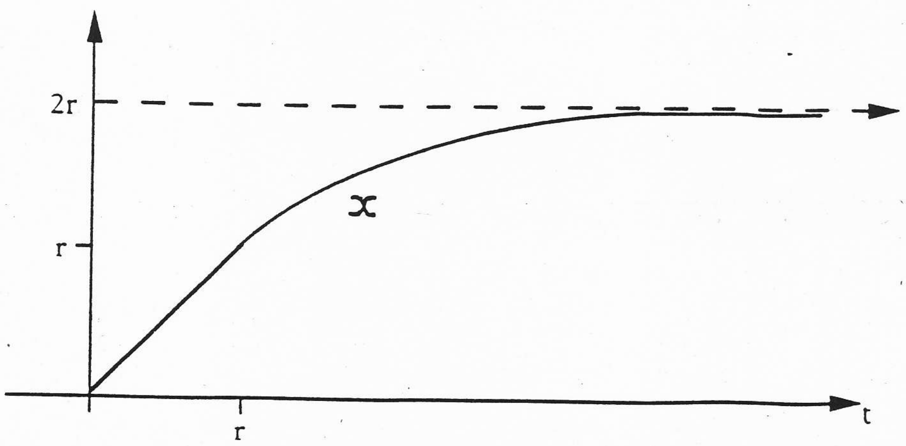

Fig. 15.

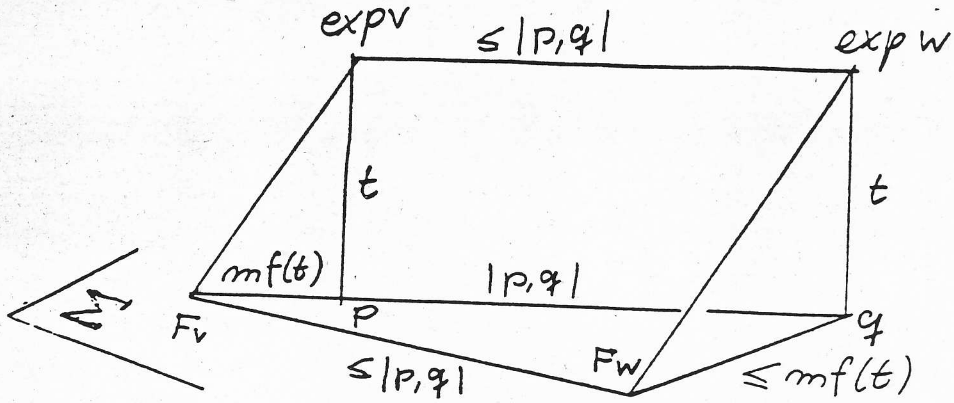

Fig. 16.

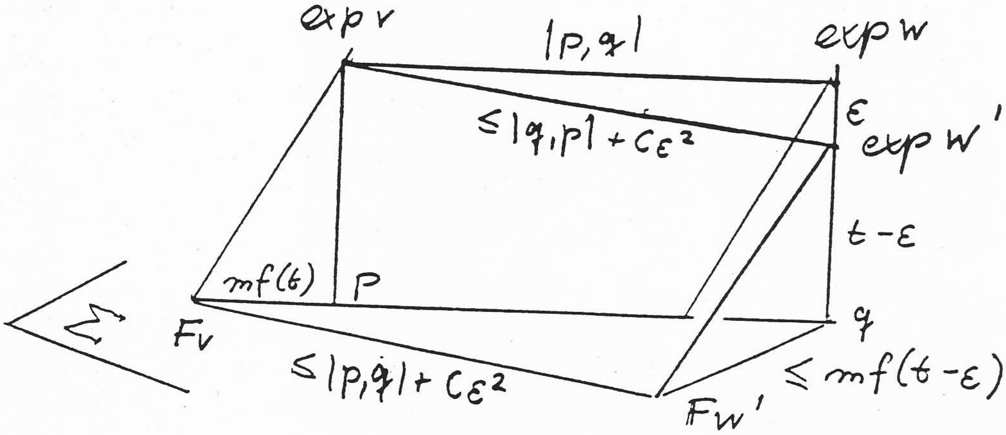

Fig. 17.

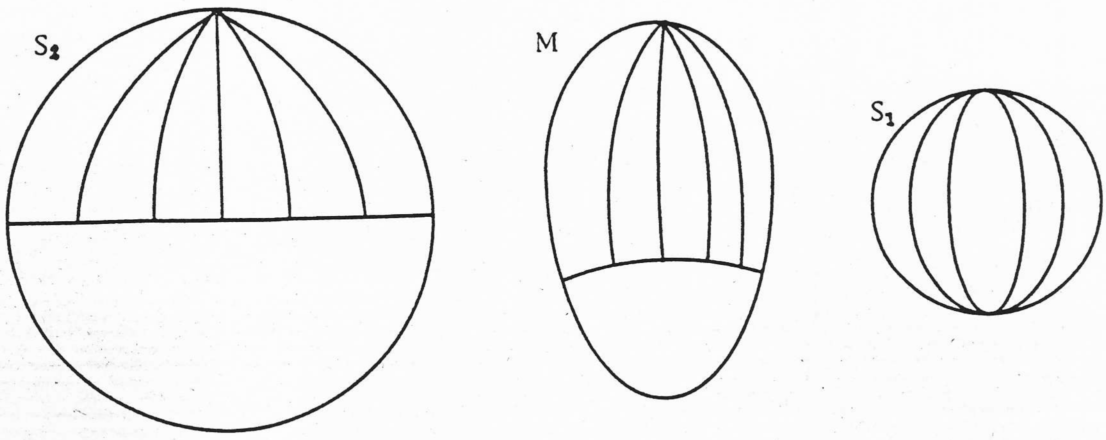

Fig. 18.

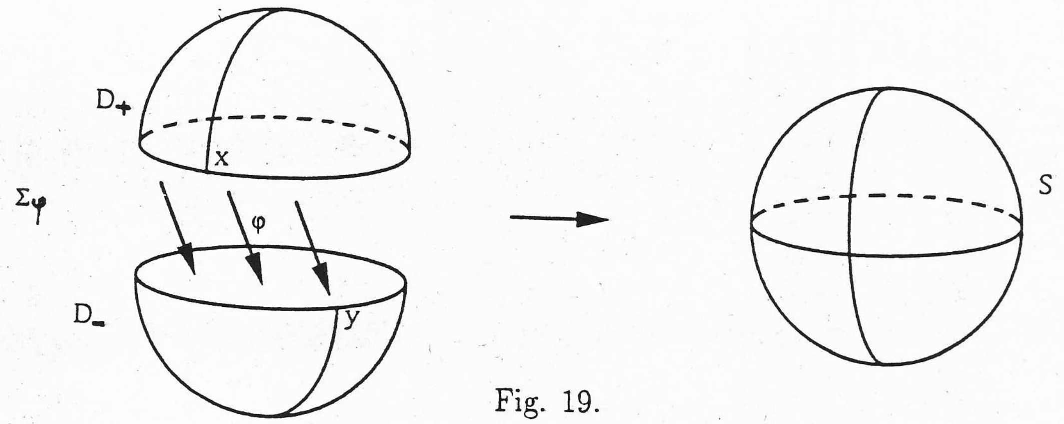

Fig. 19.

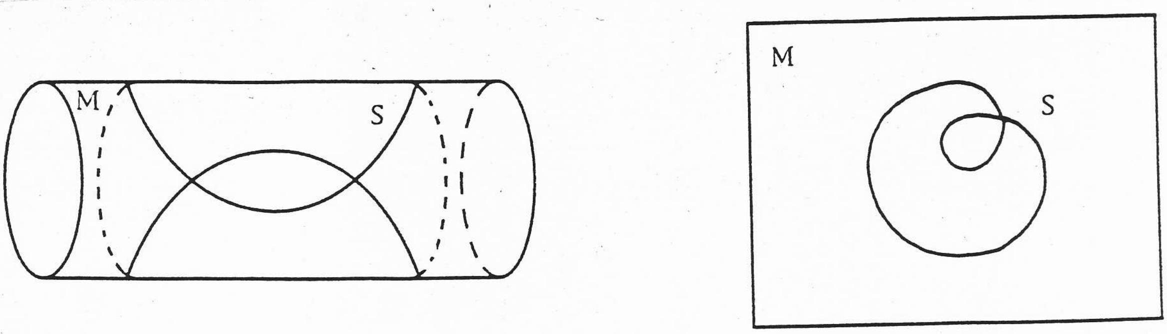

Fig. 20.

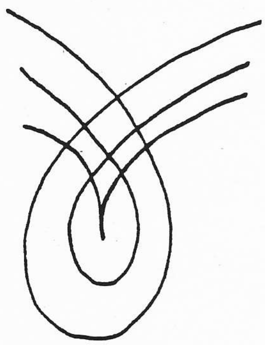

Fig. 21.

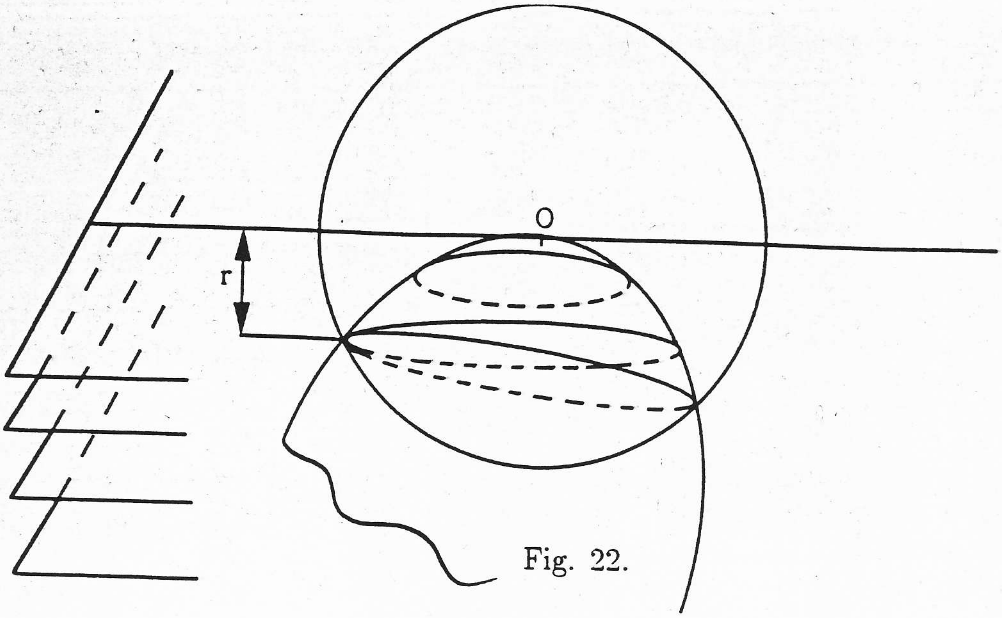

Fig. 22.

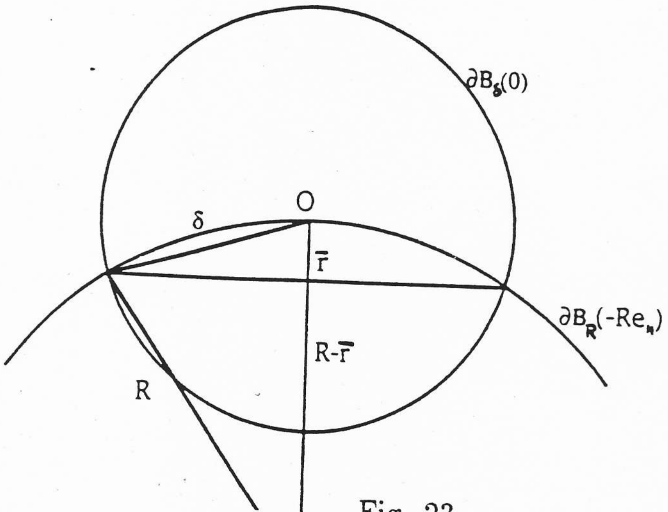

Fig. 23.

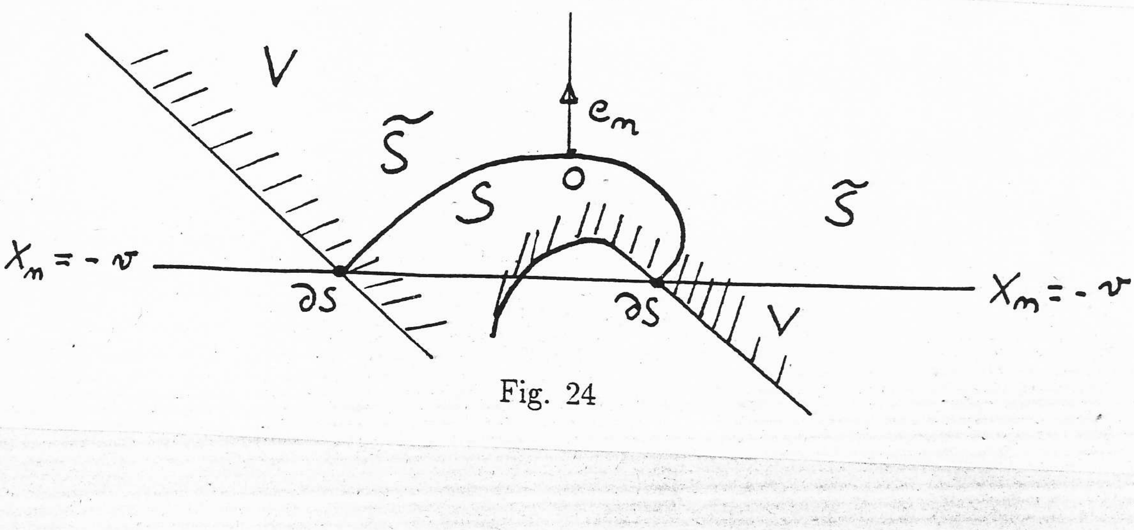

Fig. 24

second time and let $S_2 = S \setminus S_1$. All the points above S_1 are in V since they are closer to S_1 than to $\{x_n = -r\}$. Moreover, all points above S_2 are in V for the same reason.

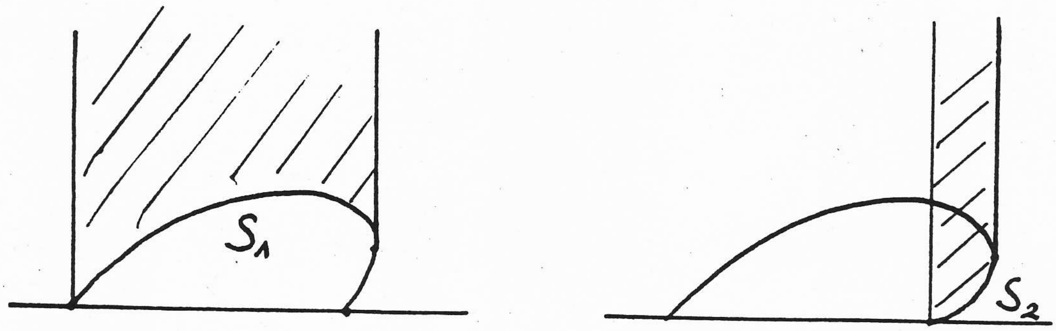

Fig. 25.

Thus we may connect any point of S to some point in $\mathbb{R}_+ \cdot e_n$ within the shaded region (cf. fig. 26 below); we just have to avoid the cylinder of height r above ∂S if we start from S_2. This finishes the proof of the claim, of the lemma and of the theorem.

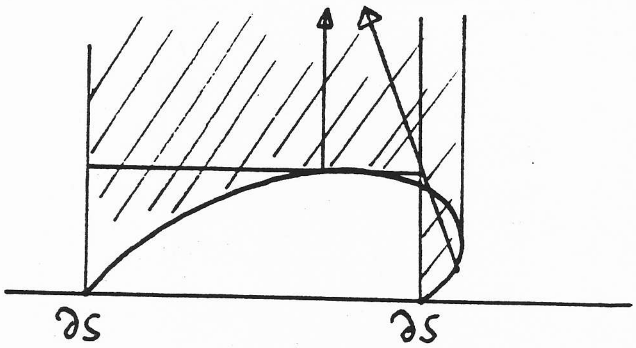

Fig. 26.

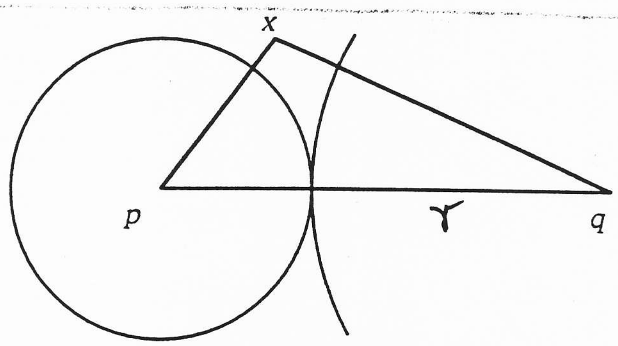

Fig. 27.

Moreover, due to $Ric \geq n-1$ and the average comparison theorem 4.1, we have on $M \setminus \{p, q\}$:

$$\Delta \rho_p \leq (n-1) \cot \rho_p, \quad \Delta \rho_q \leq (n-1) \cot \rho_q$$

in the sense of support functions. In fact, to prove the first inequality at some point $x \in M \setminus \{p, q\}$, we choose a shortest geodesic segment β from x to p and replace p by some point p' on β close to p; then the distance function $\rho_{p'}$ from p' is smooth near x and satisfies the above inequality with an arbitrary small error (by Theorem 4.1).

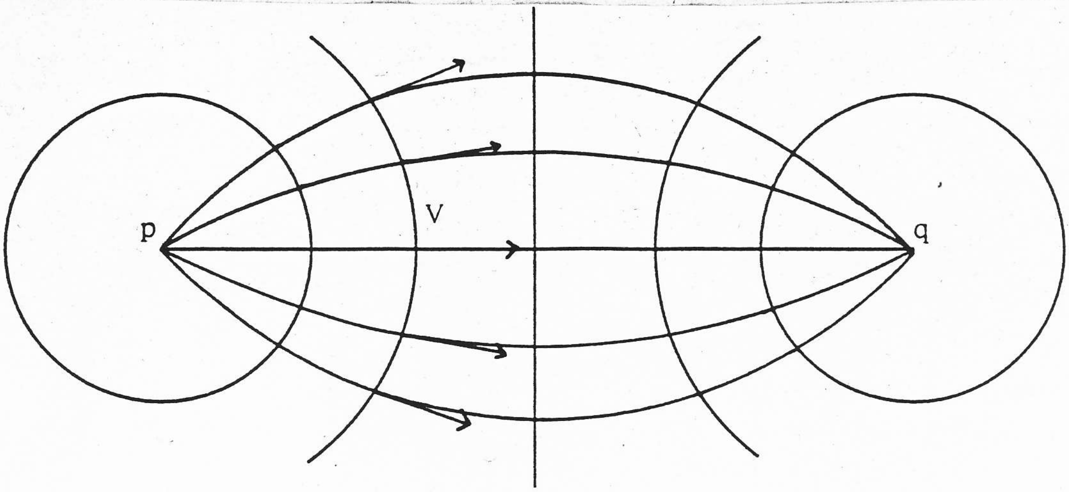

Fig. 28.