Critical Points of Distance Functions and Applications to Geometry

Jeff Cheeger

0. Introduction

These lecture notes were written for a course given at the C.I.M.E. session "Recent developments in geometric topology and related topics", June 4-12, 1990, at Montecatini Terme. Their aim is to expose three basic results in riemannian geometry, the proofs of which rely on the technique of "critical points of distance functions" used in conjunction with Toponogov's theorem on geodesic triangles. This method was pioneered by Grove and Shiohama, [GrS].

Specifically, we discuss

i) the Grove-Petersen theorem of the finiteness of homotopy types of manifolds admitting metrics with suitable bounds on diameter, volume and curvature; [GrP],

ii) Gromov's bound on the Betti numbers in terms of curvature and diameter; [G],

iii) the Abresch-Gromoll theorem on finiteness of topological type, for manifolds with nonnegative Ricci curvature, curvature bounded below and slow diameter growth; [AGl].

The first two of these theorems are stated in § 3 and proved in § 4 and §§ 5-6, respectively. The third is stated and proved in § 8.

The reader is assumed to have a background in riemannian geometry at least the rough equivalent of the first six chapters of [CE], and to be familiar with basic algebraic topology. For completeness however, the statement of Toponogov's theorem is recalled in § 2. Additional material on finiteness theorems and on Ricci curvature is provied in § 3 and § 7.

1. Critical Points of Distance Functions.

Let M^n be a complete riemannian manifold. We will assume that all geodesics are parametrized by arc length. For $p \in M^n$, we denote the distance from x to p by $\overline{x,p}$ and put

$$\rho_p(x) := \overline{x,p}$$

Note that $\rho_p(x)$ is smooth on $M \setminus \{p \cup C_p\}$, where C_p, the *cut locus* of p, is a closed nowhere dense set of measure zero.

Grove and Shiohama made the fundamental observation that there is a meaningful definition of "critical point" for such distance functions, such that in the absence of critical points, the Isotopy Lemma of Morse Theory holds. They also observed that in the presence of a lower curvature bound, Toponogov's theorem can be used to derive geometric information, from the *existence* of critical points. They used these ideas to give a short proof of a generalized Sphere Theorem, see Theorem 2.5. Other important applications are discussed in subsequent sections.

Remark 1.1. If the *sectional curvature* satisfies $K_M \le K$ (for $K \ge 0$) and q is a critical point of ρ_p with $\rho_p(q) \le \frac{\pi}{2\sqrt{K}}$, then there is also a reasonable notion of *index* which predicts the change in the *topology* when crossing a critical level. But so far, this fact has not had strong applications.

Definition 1.2. The point q $(\neq p)$ is a *critical point* of ρ_p if for all v in the tangent space, M_q, there is a minimal geodesic, γ, from q to p, making an angle, $\measuredangle\,(v, \gamma'(0)) \le \frac{\pi}{2}$, with $\gamma'(0)$. Also, p is a critical point of ρ_p.

From now on we just say that q is a critical point of p.

Remark 1.3. If $q \neq p$ is a critical point of p, then $q \in C_p$. If q is *not* critical, the collection of tangent vectors to all geodesics, γ, as above, lies in some *open* half space in M_q. Thus, there exists $w \in M_q$, such that $\measuredangle\,(w, \gamma'(0)) < \frac{\pi}{2}$, for all minimal γ from p to q.

Put $B_r(p) = \{x \mid \overline{x,p} < r\}$.

Isotopy Lemma 1.4. *If $r_1 < r_2 \le \infty$, and if $\overline{B_{r_2}(p)} \setminus B_{r_1}(p)$ is free of critical points of ρ_p, then this region is homeomorphic to $\partial B_{r_1}(p) \times [r_1, r_2]$.Moreover, $\partial B_{r_1}(p)$ is a topological submanifold (with empty boundary).*

Proof: If x is noncritical, then there exists $w \in M_x$ with $\measuredangle\,(\gamma'(0), w) < \frac{\pi}{2}$, for all minimal γ from x to p. By continuity, there exists an extension of w to a vector field, W_x, on a neighborhood, U_x, of x, such that if $y \in U_x$ and σ is minimal from y to p, then $\measuredangle\,(\sigma'(0), W_x(y)) < \frac{\pi}{2}$. Take a finite open cover of $\overline{B}_{r_2}(\rho) \setminus B_{r_1}(\rho)$, by sets, U_{x_i}, locally finite if $r_2 = \infty$, and a smooth partition of unity, $\sum \phi_i \equiv 1$, subordinate to it. Put $W = \sum \phi_i W_{x_i}$. Clearly, W is nonvanishing. For each integral curve ψ of W, the *first variation formula* gives

$$\rho_p(\psi(t_2)) - \rho_p(\psi(t_1)) \le (t_1 - t_2)\cos(\frac{\pi}{2} - \epsilon)\ ,$$

for some small ϵ. This holds on compact subsets if $r_2 = \infty$. The first statement easily follows.

To see that $\partial B_{r_1}(p)$ is a submanifold, let $q \in \partial B_{r_1}(p)$, σ a minimal geodesic from q to p, and V a small piece of the totally geodesic hypersurface at q, normal to σ. Then for $z \in V$, sufficiently close to q, each integral curve, ψ, of W through z intersects $\partial B_{r_1}(p)$ in exactly one point, $z' \in \partial B_{r_1}(p)$ (ψ extends on both sides of V). It is easy to check that the map, $z \to z'$, provides a local chart for $\partial B_{r_1}(p)$ at q.

Example 1.5. M compact and q a farthest point from p implies that q is a critical point of ρ_p, obviously, the topology changes when we pass q. This observation was made by Berger, well in advance of the formal definition of "critical point"; [Be].

Example 1.6. If γ is a geodesic loop of length ℓ and if $\gamma \mid [0, \frac{\ell}{2}]$ and $\gamma \mid [\frac{\ell}{2}, \ell]$ are minimal, then $\gamma(\frac{\ell}{2})$ is a critical point of $\gamma(0)$. In particular, if q is a closest point, to p on C_p, and q is not conjugate to p along some minimal geodesic then q is a critical point of p; see Chapter 5 of[CE]. Thus, if p, q realize the shortest distance from a point to its cut locus in M^n, and are not conjugate along any minimal γ, then p and q are *mutually critical.*

Example 1.7. On a flat torus with fundamental domain a rectangle, the barycenters of the sides and the corners project to the three critical points of p, other than p itself.

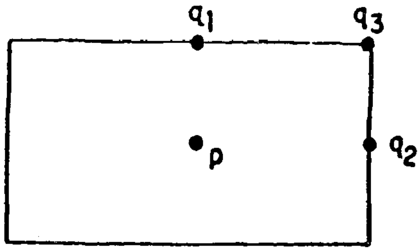

Fig. 1.1

Example 1.8. A conjugate point need not be critical. Here is a concrete example. Write the standard metric on S^2 in the form $g = dr^2 + \sin^2 r\, d\theta^2$, where $0 \le r \le \pi$, $0 \le \theta \le 2\pi$. Let $f(r, \theta)$ be a smooth function, periodic in θ, such that

i) $f(r,\theta) \equiv 1$, for all (r,θ) satisfying any of the following conditions:

$$0 \le r \le \frac{\pi}{4}\ , \qquad \frac{3}{4}\pi \le r \le \pi\ ,$$
$$\pi - \epsilon \le \theta \le \pi + \epsilon\ .$$

Here we require $\epsilon < \pi/4$.

ii) $f > 1$ elsewhere.

The metric $g' = f\,dr^2 + \sin^2 r\,d\theta^2$ satisfies $g' \geq g$. In fact, if the intersection of a curve, c, with the region, $\pi/4 < r < 3\pi/4$, is not contained in the region $\pi - \epsilon \leq \theta \leq \pi + \epsilon$, then its length with respect to g' is strictly longer than with respect to g. It follows that for the metric g', the only minimal geodesics connecting the "south pole" $(\theta = \theta)$ to the "north pole", $(\theta = \pi)$ are the curves $c(t) = (t, \theta_0)$, $\pi - \epsilon \leq \theta_0 \leq \pi + \epsilon$. Since $2\epsilon < \pi$, it follows that the north and south poles are mutually *conjugate*, but mutually *noncritical.*

We are indebted to D. Gromoll for helpful discussions concerning this example.

Remark 1.9. The *criticality radius*, r_p, is, by definition, the largest r such that $B_r(p)$ is free of critical points. By the Isotopy Lemma 1.4, $B_{r_p}(p)$ is homeomorphic to a standard open ball, since it is homeomorphic to an arbitrarily small open ball with center p.

2. Toponogov's Theorem; first applications.

Denote the length of γ by $L[\gamma]$.

By definition, a *geodesic triangle* consists of three geodesic segments, γ_i, of length $L[\gamma_i] = \ell_i$, which satisfy

$$\gamma_i(\ell_i) = \gamma_{i+1}(0) \bmod 3 \quad (i = 0, 1, 2)\ .$$

The *angle* at a corner, say $\gamma_0(0)$, is by definition, $\measuredangle(-\gamma_2'(\ell_2), \gamma_0'(0))$. The angle opposite γ_i will be denoted by α_i.

A pair of sides e.g. γ_2, γ_0 are said to determine a *hinge.*

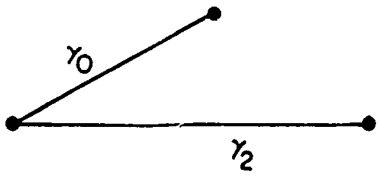

Fig. 2.1

Let M^n_H denote the n-dimensional, simply connected space of curvature $\equiv H$ (i.e. hyperbolic space, Euclidean space, or a sphere).

Toponogov's theorem has two statements. These are equivalent in the sense that either one can easily be obtained from the other.

Theorem 2.1 (Toponogov). *Let M^n be complete with curvature $K_M \geq H$.*

A) *Let $\{\gamma_0, \gamma_1, \gamma_2\}$ determine a triangle in M^n. Assume γ_1, γ_2 are minimal and $\ell_1 + \ell_2 \geq \ell_0$. If $H > 0$, assume $L[\gamma_0] \leq \dfrac{\pi}{\sqrt{H}}$. Then there is a triangle $\{\underline{\gamma}_0, \underline{\gamma}_1, \underline{\gamma}_2\}$ in M^2_H, with $L[\gamma_i] = L[\underline{\gamma}_i]$ and $\underline{\alpha}_1 \leq \alpha_1$, $\underline{\alpha}_2 \leq \alpha_2$.*

B) *Let* $\{\gamma_2, \gamma_0\}$ *determine a hinge in* M^n *with angle* α. *Assume* γ_2 *is minimal and if* $H > 0$, $L[\gamma_0] \le \frac{\pi}{\sqrt{H}}$. *Let* $\{\underline{\gamma}_2, \underline{\gamma}_0\}$ *determine a hinge in* M^2_H *with* $L[\gamma_i] = L[\underline{\gamma}_i]$, $i = 0, 2$, *and the same angle* α. *Then*

$$\overline{\gamma_2(0), \gamma_0(\ell_0)} \le \overline{\underline{\gamma}_2(0), \gamma_0(\ell_0)} .$$

Proof: See [CE], Chapter 2.

Remark 2.2. In the applications of Toponogov which occur in the sequel, the following elementary fact is often used without explicit mention. Consider the collection of hinges, $\{\underline{\gamma}_0, \underline{\gamma}_2\}$ in M^2_H, with fixed side lengths, ℓ_0, ℓ_2 and variable angle α; $0 \le \alpha \le \pi$. *Then* $\overline{\underline{\gamma}_0(\ell_0), \underline{\gamma}_2(0)}$ *is a strictly increasing function of* α.

Remark 2.3. If the inequalities in A) or B) are all equalities, more can be said (see [CE]).

By using Toponogov's theorem we can derive geometric information from the existence of critical points.

Let the triangle, $\{\gamma_0, \gamma_1, \gamma_2\}$ satisfy the hypothesis of Toponogov's theorem, and assume $\gamma_0(\ell_0)$ is *critical* with respect to $\gamma_0(0)$. Then (as explained in detail in the applications), we can

i) bound from above the side length ℓ_2 (see Theorems 2.5, 4.2),

ii) bound from below, the excess, $\ell_0 + \ell_1 - \ell_2$ (see Proposition 8.5),

iii) bound from below, the angle α_1 (see Lemma 2.6, Corollaries 2.7, 2.9, 2.10 and 6.3).

Remark 2.4. It is important to realize that in order to obtain the preceding bounds, we do *not* assume $\alpha_2 \le \pi/2$. The assumption that $\gamma_1(\ell_0)$ is *critical* with respect to $\gamma_0(0)$ implies that $\sphericalangle(-\tilde{\gamma}_0'(\ell_0), \gamma_1'(0)) \le \pi/2$, for *some* minimal $\tilde{\gamma}_0$ from $\gamma_0(0)$ to $\gamma_0(\ell_0)$. This is all that we require.

Theorem 2.5 (Grove-Shiohama). *Let* M^n *be complete*, with $K_M \ge H$, *for some* $H > 0$. *If* M^n *has diameter*, $\mathrm{dia}(M^n) > \frac{\pi}{2\sqrt{H}}$, *then* M^n *is homeomorphic to the sphere*, S^n.

Proof: Let $p, q \in M^n$ be such that $\overline{p,q} = \mathrm{dia}(M^n)$; in particular, p and q are mutually critical (see Example 1.5).

Claim. There exists no $x \ne q, p$ which is critical with respect to p (the same holds for q).

Proof of Claim: Assume x is such a point. Let γ_2 be minimal from q to x. By assumption there exists γ_0, minimal from x to p, with

$$\alpha_1 = \sphericalangle(-\gamma_2'(\ell_2), \gamma_0'(0)) \le \frac{\pi}{2} .$$

Similarly, since p and q are mutually critical, there exist *minimal* γ_1, $\tilde{\gamma}_1$ from p to q such that

$$\sphericalangle(-\gamma_0'(\ell_0), \gamma_1'(0)) \le \frac{\pi}{2}$$

and

$$\sphericalangle(-\tilde{\gamma}_1'(\ell_1), \gamma_2'(0)) \le \frac{\pi}{2} .$$

Note that $L[\gamma_1] = L[\tilde{\gamma}_1] = \overline{p,q} > \frac{\pi}{2\sqrt{H}}$.

Apply A) of Toponogov's theorem to both $\{\gamma_0, \gamma_1, \gamma_2\}$ and $\{\gamma_0, \tilde{\gamma}_1, \gamma_2\}$. Since a triangle in M_H^2 (the sphere) is determined up to congruence by its side lengths, we get a *unique* triangle, $\{\underline{\gamma}_0, \underline{\gamma}_1, \underline{\gamma}_2\}$, in M_H^2, *all* of whose angles are $\leq \pi/2$. By elementary spherical trigonometry, this implies that all sides have length $\leq \dfrac{\pi}{2\sqrt{H}}$, contradicting $\overline{p, q} > \dfrac{\pi}{2\sqrt{H}}$.

Given the claim, the proof is easily completed (compare the proof of Reeb's Theorem given in [M]).

The following observation and its corollaries (2.7, 2.10) are of great importance.

Lemma 2.6 (Gromov). *Let q_1 be critical with respect to p and let q_2 satisfy*

$$\overline{p, q_2} \geq \nu \overline{p, q_1} \ ,$$

for some $\nu > 1$. Let γ_1, γ_2 be minimal geodesics from p to q_1, q_2 respectively and put $\theta = \measuredangle\,(\gamma_1'(0), \gamma_2'(0))$.

i) *If $K_M \geq 0$,*

$$\theta \geq \cos^{-1}(1/\nu) \ .$$

ii) *If $K_M \geq H$, $(H < 0)$ and $\overline{p, q_2} \leq d$, then*

$$\theta \geq \cos^{-1}\left(\frac{\tanh(\sqrt{-H}d/\nu)}{\tanh(\sqrt{-H}d)}\right) \ .$$

Proof: Put $\overline{p, q_1} = x$, $\overline{q_1, q_2} = y$, $\overline{p, q_2} = z$. Let σ be minimal from q_1 to q_2. Since q is critical for p, there exists τ, minimal from q to p with

$$\measuredangle\,(\sigma'(0), \tau'(0)) \leq \frac{\pi}{2} \ .$$

i) Applying Toponogov's Theorem B) to the hinges $\{\sigma, \tau\}$ and $\{\gamma_1, \gamma_2\}$ gives

$$z^2 \leq x^2 + y^2 \ ,$$

$$y^2 \leq x^2 + z^2 - 2xz\cos\theta \quad \text{(law of cosines)}$$

Since $z \geq \nu \cdot x$, the conclusion easily follows.

ii) By scaling, we can assume $H = -1$. Replace the inequalities above by the following ones from hyperbolic trigonometry (see e.g. [Be])

$$\cosh z \leq \cosh x \cosh y \ ,$$

$$\cosh y \leq \cosh x \cosh z - \sinh x \sinh z \cos\theta \ .$$

Substituting the second of these into the first and simplifying gives

$$\theta \geq \cos^{-1}\left(\frac{\tanh x}{\tanh z}\right) ,$$

which suffices to complete the proof.

Corollary 2.7. *Let $q_1, \ldots, q_N$ be a sequence of critical points of p, with*

$$\overline{p, q_{i+1}} \geq \nu \, \overline{p, q_i} \quad (\nu > 1)$$

i) *If $K_{M^n} \geq 0$ then*

$$N \leq \mathcal{N}(n, \nu)$$

ii) *If $K_M \geq H$ $(H < 0)$ and $q_N \leq d$, then*

$$N \leq \mathcal{N}(n, \nu, Hd^2) \ .$$

Proof: Take minimal geodesics, γ_i from p to q_i. View $\{\gamma_i'(0)\}$ as a subset of $S^{n-1} \subset M_p^n$. Then Lemma 2.6 gives a lower bound on the distance, θ, between any pair $\gamma_i'(0), \gamma_j'(0)$. The balls of radius $\theta/2$ about the $\gamma_i'(0) \in S^{n-1}$ are mutually disjoint. Hence, if we denote by $V_{n-1,1}(r)$, the volume of a ball of radius r on S^{n-1}, we can take

$$\mathcal{N} = \frac{V_{n-1,1}(\pi)}{V_{n-1,1}(\theta/2)} \ ,$$

where $V_{n-1,1}(\pi) = \mathrm{Vol}(S^{n-1})$ and θ is the minimum value allowed by Lemma 2.6.

Remark 2.8. It turns out that Corollary 2.7 is the only place in which the hypothesis on *sectional* curvature is used in deriving Gromov's bound on Betti numbers in terms of curvature and diameter. For details, see Theorem 3.8 and §§ 5-6.

The following result is a weak version (with a much shorter proof) of the main result of [CG12], compare also § 8.

Corollary 2.9. *Let M^n be complete,* with $K_{M^n} \geq 0$. *Given p, there exists a compact set C, such that p has no critical points lying outside C. In particular M^n is homeomorphic to the interior of a compact manifold with boundary.*

Proof: The first statement, which is obvious from Corollary 2.7, easily implies the second.

Corollary 2.10. *Let $\mathcal{N}(n, \nu, Hd^2)$ be as in* Corollary 2.7, *and let $r_1 \nu^{\mathcal{N}} < r_2$. Then there exists $(s_1, s_2) \subset [r_1, r_2]$ such that $\rho_p^{-1}((s_1, s_2))$ is free of critical points and*

$$s_2 - s_1 \geq (r_2 - r_1 \nu^{\mathcal{N}})(1 + \nu + \cdots \nu^{\mathcal{N}})^{-1}$$

Moreover, the set of critical points has measure at most $(1 - \nu^{-\mathcal{N}}) r_2$.

Proof: Let $r_1 + \ell_1$ denote the first critical value $\geq r_1$; $\ell_2 + \nu(r_1 + \ell_1)$ the first after $\nu(r_1 + \ell_1)$ etc. It is easy to see that in the worst case

$$\ell_1 = \ell_2 = \cdots = \ell \ ,$$

$$(\cdots(\nu(\nu(r_1 + \ell) + \ell) + \ell \cdots) + \ell = r_2$$

The first assertion follows easily. The proof of the second is similar.

Remark 2.11. The proof of Corollary 2.7 easily yields an explicit estimate for the constant $\mathcal{N}$. For example, in case $K_{M^n} \geq 0$, we get

$$\mathcal{N}(n,\nu) \leq \left(\frac{\pi}{\frac{1}{2}\cos^{-1}(1/\nu)}\right)^{n-1}$$

Thus, for ν close to 1,

$$\mathcal{N}(n,\nu) \leq \left[\frac{2\pi^2}{(\nu-1)}\right]^{(n-1)/2}$$

3. Background on Finiteness Theorems.

The theorems in question bound topology in terms of bounds on geometry. In subsequent lectures we will prove two such results due to Gromov, [G] and Grove-Petersen [GrP]. Before stating these, we establish the context by giving an earlier result of Cheeger [C1], [C3] (see also [GLP], [GreWu], [Pe1], [Pe2], [We] for related developments).

Theorem 3.1. (Cheeger). *Given $n, d, V, K > 0$, the collection of compact n-manifolds which admit metrics whose diameter, volume and curvature satisfy,*

$$\begin{aligned} \operatorname{dia}(M^n) &\leq d\ , \\ \operatorname{Vol}(M^n) &\geq V\ , \\ |K_M| &\leq K\ , \end{aligned}$$

contains only a finite number, $C(n,,V^{-1}d^n, Kd^2)$, of diffeomorphism types.

Remark 3.2. The basic point in the proof is to establish a lower bound on the length of a smooth closed geodesic (here one need only assume $K_M \geq K$). This, together with the assumption $K_M \leq K$, gives a lower bound on the injectivity radius of the exponential map (see [CE], Chapter 5). Although Theorem 3.1 predated the use of critical points, the crucial ingredient in the Grove-Petersen theorem below is essentially a generalization of the above mentioned lemma on closed geodesics (compare Example 1.6).

Theorem 3.3 (Grove-Petersen). *Given $n, d, V > 0$ and H, the collection of compact n-dimensional manifolds which admit metrics satisfying*

$$\begin{aligned} \operatorname{dia}(M^n) &\leq d\ , \\ \operatorname{Vol}(M^n) &\geq V\ , \\ K_M &\geq H\ , \end{aligned}$$

contains only a finite number, $C(n, V^{-1}d^n, Hd^2)$, of homotopy types.

Remark 3.4. In [GrPW], the conclusion of Theorem 3.3 is strengthened to finiteness up to homeomorphism ($n \neq 3$) and up to diffeomorphism ($n \neq 3,4$). The proof employs techniques from "controlled topology". Thus, Theorem 3.3 supersedes Theorem 3.1 (as stated) if $n \neq 3,4$. However, Theorem 3.1 can actually be strengthened to give a conclusion which does not hold under the hypotheses of Theorem 3.3.

Given $\{M_i^n\}$ as in Theorem 3.1, there is a subsequence $\{M_j^n\}$, a manifold M^n, and diffeomorphisms, $\phi_j : M^n \to M_j^n$, such that the pulled back metrics, $\phi_j^*(g_j)$, converge in the $C^{1,\alpha}$-topology, for all $\alpha > 1$ (see the references given at the beginning of this section for further details).

Example 3.5. By rounding off the tip of a cone, a surface of nonnegative curvature is obtained. From this example, one sees that under the conditions of Theorem 3.3, arbitrarily small metric balls need not be contractible. Thus, the criticality radius can be arbitrarily small (compare Remark 1.9).

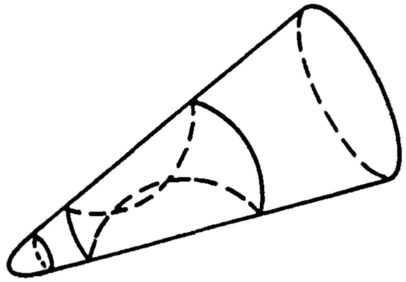

Fig. 3.1

However, it will be shown that the inclusion of a sufficiently small ball into a somewhat larger one is homotopically trivial.

Example 3.6. Consider the *surface* of a solid cylindrical block from which a large number, j, of cylinders (with radii tending to 0) have been removed.

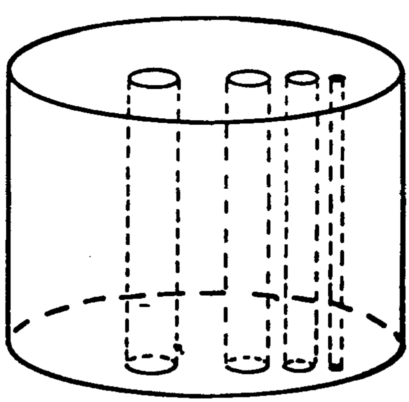

Fig. 3.2

The edges can be rounded so as to obtain a manifold, M_j^2, with $\mathrm{Vol}(M_j^2) \geq V$, $\mathrm{dia}(M_j^2) \leq d$ (but inf $K_{M_j^2} \to -\infty$, as $j \to \infty$). For the first Betti number, one has $b^1(M_j^2) = 2j \to \infty$.

Note that the metrics in this sequence can be rescaled so that $K_{M_j^2} \geq -1$, $\mathrm{Vol}(M_j^2) \to \infty$. Then, of course, $\mathrm{dia}(M_j^2) \to \infty$ as well.

Example 3.7. Consider the lens space L_n^3, obtained by dividing

$$S^3 = \{(z_1, z_2) \mid |z_1|^2 + |z_2|^2 = 1\} ,$$

by the action of $\mathbb{Z}_n = \{1, a, \ldots, a^{n-1}\}$, where $a : (z_1, z_2) \to (e^{2\pi i/n} z_1, e^{2\pi i/n} z_2)$. Then $\mathrm{dia}(L_n^3) = 1$, $K_{L_n^3} \equiv 1$, but $\mathrm{Vol}(M_n^3) \to 0$, and $H_1(L_n^3, \mathbb{Z}) = \mathbb{Z}_n$. Thus, if the lower bound on volume is relaxed, there are infinitely many possibilities for the first homology group, H_1. Nonetheless, the following theorem of Gromov asserts that for any *fixed* coefficient field, F, the Betti numbers, $b^i(M^n)$ are bounded independent of F.

Theorem 3.8 (Gromov). *Given $n, d > 0$, H, and a field F,* if

$$\mathrm{dia}(M^n) \leq d ,$$

$$K_{M^n} \geq H ,$$

then

$$\sum_i b^i(M^n) \leq C(n, Hd^2) .$$

Corollary 3.9. *If M^n has nonnegative sectional curvature, $K_{M^n} \geq 0$, then*

$$\sum_i b^i(M^n) \leq C(n) .$$

Remark 3.10. The most optimistic conjecture is that $K_{M^n} \geq 0$ implies $b^i(M^n) \leq \binom{n}{i}$, and hence, $\sum_i b^i(M^n) \leq 2^n$. Note $b^i(T^n) = \binom{n}{i}$ where T^n is a flat n-torus. At present, one knows only that $K_{M^n} \geq 0$ (in fact $\mathrm{Ric}_{M^n} \geq 0$) implies $b^1(M^n) \leq n$. But the method of proof of Theorem 3.8 does not give this sharp estimate; compare also [GLP], p. 72.

In proving Theorems 3.1, 3.3 and 3.8, a crucial point is to bound the number of balls of radius ϵ needed to cover a ball of radius r.

Proposition 3.11 (Gromov). *Let the Ricci curvature of M^n satisfy $\mathrm{Ric}_{M^n} \geq (n-1)H$. Then given $r, \epsilon > 0$ and $p \in M^n$, there exists a covering, $B_r(p) \subset \cup_1^N B_\epsilon(p_i)$, $(p_i \in B_r(p))$ with $N \leq N_1(n, Hr^2, r/\epsilon)$. Moreover, the multiplicity of this covering is at most $N_2(n, Hr^2)$.*

Remark 3.12. The condition $\mathrm{Ric}_{M^n} \geq (n-1)H$ is implied by $K_{M^n} \geq H$, in which case, the bound on N_1 could be obtained from Toponogov's theorem. For the proof of Proposition 3.11, see § 7.

Remark 3.13. The conclusion of Theorem 3.8 (and hence of Corollary 2.7) fails if the hypothesis $K_m \geq H$ is weakened to the lower bound on Ricci curvature, $Ric_{M^n} \geq (n-1)H$; see [An], [ShY].

Remark 3.14. S. Zhu has shown that homotopy finiteness continues to hold for $n = 3$, if the lower bound on sectional curvature is replaced by a lower bound on Ricci curvature; [Z]. Whether or not this remains true in higher dimensions is an open problem.

4. Homotopy finiteness.

Pairs of mutually critical points.

The main point in proving the theorem on homotopy finiteness is to establish a lower bound on the distance between a pair of *mutually critical* points (compare Example 1.6). For technical reasons we actually need a quantitative refinement of the notion of criticality.

Definition 4.1. q is *ϵ-almost critical* with respect to p, if for all $v \in M_q$, there exists γ, minimal from q to p, with $\sphericalangle\,(v,\gamma'(0)) \leq \frac{\pi}{2} + \epsilon$.

Theorem 4.2. *There exist $\epsilon = \epsilon(n, V^{-1}d^n, Hd^2)$, $\delta = \delta(n, V^{-1}d^n, Hd^2) > 0$, such that if $p, q \in M^n$*

$$\mathrm{dia}(M^n) \leq d \ ,$$
$$\mathrm{Vol}(M^n) \geq V \ ,$$
$$K_{M^n} \geq H \ ,$$
$$\overline{p,q} < \delta d \ ,$$

then at least one of p, q is not ϵ-almost critical with respect to the other.

The proof of Theorem 4.2 uses two results on volume comparison. The first of these, Lemma 4.3, is stated below and proved in the Appendix to this section. The second result, Proposition 4.7 is stated and proved in the Appendix.

For $X \subset Y$ closed, put

$$T_r(X) = \{q \in Y \mid \overline{q, X} < r\}$$

(the case of interest below is $Y = S^{n-1}$, the unit $(n-1)$-sphere).

Recall that the volume of a ball in M^n_H is given as follows. Put

$$\mathcal{A}_{n-1,H}(s) = \begin{cases} (\frac{1}{\sqrt{H}} \sin s\sqrt{H}\, s)^{n-1} & H > 0 \\ s^{n-1} & H = 0 \\ (\frac{1}{\sqrt{-H}} \sinh \sqrt{-H}\, s)^{n-1} & H < 0 \end{cases}$$
$$V_{n,H}(r) = v_{n-1} \int_0^r \mathcal{A}_{n-1,H}(s)\, ds \ ,$$

where $v_{n-1} = V_{n-1,1}(\pi)$ is the volume of the unit $(n-1)$-sphere. Then in M^n_H,

$$\mathrm{Vol}(B_r(\underline{p})) = V_{n,H}(r) \ .$$

Lemma 4.3. *Let $X \subset S^n$ be closed. Then*

a)
$$\frac{\mathrm{Vol}(T_{r_1}(X))}{\mathrm{Vol}(T_{r_2}(X))} \geq \frac{V_{n,1}(r_1)}{V_{n,1}(r_2)} \ .$$

Thus,

b)
$$\frac{\mathrm{Vol}(T_{r_2}(X)) - \mathrm{Vol}(T_{r_1}(X))}{\mathrm{Vol}(T_{r_2}(X))} \le \frac{V_{n,1}(r_2) - V_{n,1}(r_1)}{V_{n,1}(r_2)} .$$

Remark 4.4. The lemma actually holds for $X \subset M^n$, where $\mathrm{Ric}_{M^n} \ge (n-1)H$ provided $V_{n,1}(r)$ is replaced by $V_{n,H}$ (see Proposition 7.1).

Proof of Theorem 4.2: By scaling, we can assume $d = 1$.

In i)-iii) below we determine ϵ, δ. In iv) we show that they have the desired properties.

i) Fix $\epsilon > 0$. Let $\underline{\gamma}_0, \underline{\gamma}_1$ determine a hinge in M_H^2 with angle,

$$\alpha < \frac{\pi}{2} - \epsilon ,$$

at the point,

$$\underline{\gamma}_0(\ell_0) = \underline{\gamma}_1(0)$$

Here, $L[\underline{\gamma}_i] = \ell_i$. Let $\delta = \delta(H, \epsilon, r)$ be the length of the base of an isosceles triangle in M_H^2 with equal sides of length r and angle $\pi/2 - \epsilon$ opposite these sides. Then if

$$\ell_0 \le \delta ,$$

$$\ell_1 \ge r ,$$

we have

$$\overline{\underline{\gamma}_0(0), \underline{\gamma}_1(\ell_1)} < \ell_1 .$$

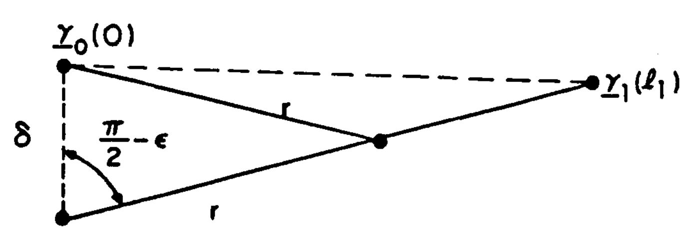

Fig. 4.1

ii) Determine $\epsilon = \epsilon(n, V^{-1}, H)$ by

$$\frac{V_{n-1,1}(\frac{\pi}{2}+\epsilon) - V_{n-1,1}(\frac{\pi}{2}-\epsilon)}{V_{n-1,1}(\frac{\pi}{2}+\epsilon)} \times V_{n,H}(1) = \frac{V}{6} .$$

iii) Determine $r = r(n, V^{-1}, H)$ by

$$V_{n,H}(r) = \frac{V}{6} .$$

iv) Assume p, q are mutually ϵ-almost critical and that $\overline{p,q} < \delta$ with δ, ϵ, r as in i)-iii). We claim $\mathrm{Vol}(M^n) \le \frac{2}{3}V$ which is a contradiction.

Let

$$M^n(p) = \{x \in M^n \mid \overline{x,p} < \overline{x,q}\}$$

Since $M \setminus (M^n(p) \cup M^n(q))$ has measure zero, by symmetry, it suffices to show

$$\mathrm{Vol}(M^n(p)) \le \frac{V}{3} .$$

Let $X \subset S^{n-1} \subset M_p^n$ be the set of tangent vectors to minimal geodesics from p to q. By assumption, $\overline{T_{(\pi/2)+\epsilon}(X)} = S^{n-1}$. Hence by ii), Lemma 4.3 b) and Proposition 4.7, the volume of the set of points, $x \in M^n \setminus C_p$, such that $x = \gamma(\ell)$, and $\gamma'(0) \notin T_{(\pi/2)-\epsilon}(X)$, is at most $V/6$.

But if $y = \sigma(u)$, $\sigma'(0) \in T_{(\pi/2)-\epsilon}(X)$ and $u > r$, then by i) and Toponogov's theorem B), we have

$$\overline{q,y} < \overline{p,y} .$$

Therefore $y \notin M^n(p)$.

By the choice of r (see iii)) the set of such $y \in M^n(p)$ has volume $\le V/6$ (see again Proposition 4.7). Thus, we get the contradiction

$$\mathrm{Vol}(M^n(p)) \le \frac{V}{6} + \frac{V}{6} = \frac{V}{3} .$$

Let $\Delta \subset M \times M$ denote the diagonal.

Corollary 4.5. *Let M^n, δ be as above. Then there exists a deformation retraction H_t : $T_{(\delta/2)d}(\Delta) \to \Delta$ ($t \in [0,1]$) such that the curves, $t \to H_t(p,q)$ have length*

$$L[H_t(p,q)] \le R(n, V^{-1}d^n, Hd^2)\,\overline{p,q} .$$

Proof: By scaling, it suffices to assume $d = 1$. Let $(p,q) \in T_{\delta/2}(\Delta)$ with say q not ϵ-critical with respect to p. Let U_q, W_q be as in the Isotopy Lemma 1.4. Let $W'_{(p,q)}$ be the vector field $(0, W_q)$ on some sufficiently small neighborhood $V_p \times W_q$. (By averaging under the flip, we can even replace $W'_{(p,q)}$ by $W''_{(p,q)}$ such that $W''_{(p,q)} = W''_{(q,p)}$). The proof is completed by a partition of unity construction and first variation argument like those in the Isotopy Lemma 1.4 (the deformation we obtain does not necessarily preserve $T_{\delta/2}(\Delta)$, but satisfies the estimate above).

Curves varying continuously with their endpoints.

Let $(p,q) \in T_{\delta/2}(\Delta)$. Write $H_t(p,q) = (\phi_1(t,p,q),\ \phi_2(t,p,q))$. Put

$$\phi(t,p,q) = \begin{cases} \phi_1(2t,p,q) & 0 \le t \le \frac{1}{2} \\ \phi_2(1-2t,p,q) & \frac{1}{2} \le t \le 1 \end{cases}$$

Then $\phi(t,p,q)$ is a curve from p to q, which depends continuously on (p,q), with

$$L[\phi(t,p,q)] \leq R\,\overline{p,q}\ .$$

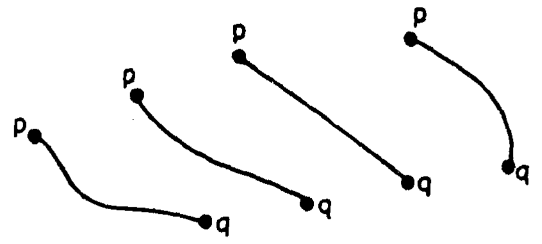

Fig. 4.2

Maps which are close are homotopic.

Corollary 4.6. *Let N be arbitrary and $f_i : N \to M^n$ (as above) $i = 1,2$. If $\overline{f_1(x), f_2(x)} < \frac{\delta}{2} d$, for all $x \in N$, then f_0, f_1 are homotopic.*

Proof: The required homotopy is given by

$$f_t(x) = \phi(t, f_0(x), f_1(x))\ .$$

Mapping in simplices (center of mass).

Let $(p_0, \ldots, p_k) \in M \times \cdots \times M$, such that $\overline{p_i, p_j} \leq (1 + \cdots + R^{k-1})^{-1}\frac{\delta}{2} d$. Construct a map of a k-simplex into M^n, inductively as follows.

i) Join p_0 to p_1 by $\phi(t,p_0,p_1)$

ii) Join p_2 to each point of $\phi(t,p_0,p_1)$ by $\phi(s,\phi(t,p_0,p_1),p_2)$.

iii) Join p_3 to each point of $\phi(s,\phi(t,p_0,p_1),p_2)$ by $\phi(u,\phi(s,\phi(t,p_0,p_1),p_2),p_3)$, etc.

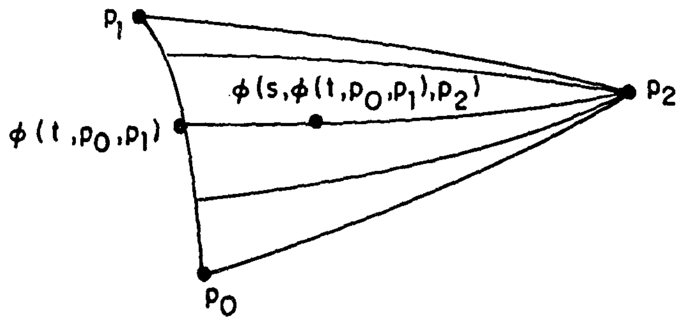

Fig. 4.3

After an obvious reparametrization, we get a map to M^n of the simplex $(\alpha_0, \dots, \alpha_k)$, $0 \leq \alpha_i \leq 1$, $\sum \alpha_i = 1$.

Proof of Theorem 3.3:

i) By scaling, the metric we can assume $d = 1$

ii) By Proposition 3.11, for any such M^n, we can fix a covering,

$$M^n = \cup B_\epsilon(p_i) \ , \ \text{where } 0 \leq i \leq N_1(n, H, \epsilon, 10) \ .$$

Moreover, the multiplicity of this covering is $\leq N_2(n, H)$. Take

$$\epsilon = \frac{\delta}{12}(1 + R + \cdots + R^{N_1 - 1})^{-1} \ ,$$

with δ, R as in Theorem 4.1.

iii) Since $0 < \overline{p_i, p_j} \leq 1$, by the "pigeon-hole" principle, we can divide the collection of all such M^n into $C(n, V^{-1}, H)$ classes, such that if M_1^n, M_2^n are in the same class,

$$\operatorname{card}(\{p_{i,1}\}) = \operatorname{card}(\{p_{i,2}\}) = c \leq N_1 \ ,$$

and for $\{p_{i,\ell}\} \subset M_\ell^n$, $\ell = 1, 2$ as above,

$$|\overline{p_{i,1}, p_{j,1}} - \overline{p_{i,2}, p_{j,2}}| \leq \frac{\delta}{12}(1 + R + \cdots + R^{N_1 - 1})^{-1} \ .$$

iv) It suffices to show that M_1^n, M_2^n as in iii) are homotopy equivalent. Construct a map $h_1 : M_1^n \to M_2^n$ as follows. Choose a partition of unity, $\sum \phi_i \equiv 1$, subordinate to $\{B_\epsilon(p_{i,1})\}$. Define a map η_1 of M_1^n to the standard $(c-1)$-simplex by

$$\eta_1(x) = (\phi_0(x), \dots, \phi_{c-1}(x)) \ .$$

Let K be the subcomplex consisting of those closed simplices whose interior has nonempty intersection with range η_1. It follows from iii) that for any $\sigma \in K$ we can define a map $g_1 : \sigma \to M_2^n$ (using the center of mass construction). The maps on the various $\sigma \subset K$ fit together to give $g_1 : K \to M_2^n$. Put $h_1 = g_1 \eta_1$ and define $h_2 : M_2^n \to M_1^n$ similarly.

Let Id_{M_j} denote the identity map on M_j. It is easy to see that the pairs $(h_2 h_1, Id_{M_1})$, $(h_1 h_2, Id_{M_2})$ satisfy the hypothesis of Corollary 4.6. q.e.d.

Appendix

Proof of Lemma 4.3: For all $\epsilon > 0$, we can find a finite set of points, X_ϵ, such that $T_\epsilon(X_\epsilon) \supset X$, $T_\epsilon(X) \supset X_\epsilon$. Since ultimately, we can let $\epsilon \to 0$, it follows easily that it suffices to assume that $X = p_1 \cup \cdots \cup p_N$ is itself a finite set. Fix i and define the *starlike* set, U_r, by

$$U_r = \{x \in T_r(X) \mid \overline{x, p_i} < \overline{x, p_j} \quad \forall j \neq i\} \ .$$

Since i is arbitrary, it clearly suffices to show

$$\frac{\mathrm{Vol}(U_{r_1})}{\mathrm{Vol}(U_{r_2})} \geq \frac{V_{n,1}(r_1)}{V_{n,1}(r_2)} .$$

Put

$$\{A_{r_1} = \gamma(s) \mid \gamma(s) \in U_{r_1} \text{ and } \exists t > r_1 \text{ with } \gamma(t) \in U_{r_2}\} .$$

Then, clearly we get

$$\frac{Vol(U_{r_1})}{Vol(U_{r_2})} \geq \frac{Vol(A_{r_1})}{Vol(U_{r_2})} \geq \frac{V_{n,1}(r_1)}{V_{n,1}(r_2)} . \quad q.e.d.$$

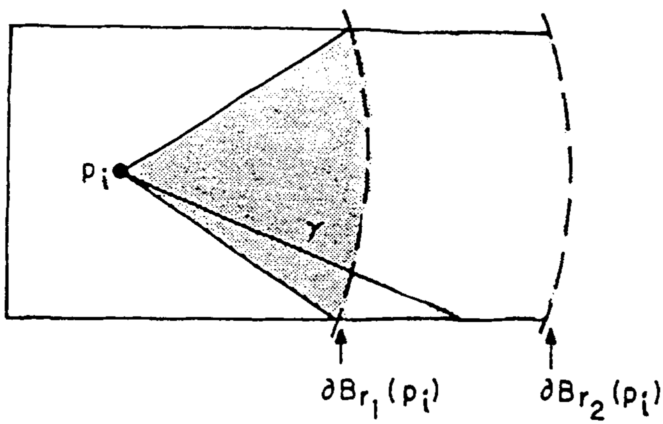

Fig. 4.4

Let γ be a geodesic with $\gamma(0) = p$ and let $\gamma(\ell_j)$ denote the cut point of γ. Let U be the interior of the cut locus in M^n_p. Thus,

$$U = \{t\gamma'(0) \mid 0 \leq t < \ell_j\} .$$

Then $\exp_p \overline{U} = M^n$.

Let $I : M^n_p \to (\underline{M}^n_H)_{\underline{p}}$ be an isometry and put

$$\underline{U} = \exp_{\underline{p}} \cdot I(U) .$$

Proposition 4.7. *The map*

$$\exp_p \cdot I^{-1} \cdot \exp_{\underline{p}}^{-1} \mid \overline{\underline{U}}$$

is distance decreasing if $K_{M^n} \geq H$ and volume decreasing if $\mathrm{Ric}_{M^n} \geq (n-1)H$.

Proof: The first assertion is clear from Toponogov's theorem B). In particular the above map is volume decreasing in this case. For the second assertion, see Remark 7.3.

5. Betti numbers and rank.

Gromov's inequality (Theorem 3.8)

$$\sum b^i(M^n) \le C(n, Hd^2)$$

depends on a novel method of estimating Betti numbers (as well as on the interaction between curvature and critical point theory, in particular, Corollary 2.7). In the present section, we estimate Betti numbers in terms of an invariant called *rank*. This part of the discussion (and much of that of §6) applies to metric spaces considerably more general than riemannian manifolds. In §6, we show that in the context of Theorem 3.8, rank can be estimated in terms of curvature and diameter (specifically in terms of the numbers $\mathcal{N}(n, Hd^2)$ of Corollary 2.7 and $N_1(n, Hd^2, 10^{n+1})$ of Proposition 3.11).

Let $U_1, U_2 \subset M$ be open. The Mayer-Vietoris sequence,

$$\to H^{i-1}(U_1 \cap U_2) \to H^i(U_1 \cup U_2) \to H^i(U_1) \oplus H^i(U_2) \to$$

leads immediately to the estimate

$$b^i(U_1 \cup U_2) \le b^i(U_1) + b^i(U_2) + b^{i-1}(U_1 \cap U_2)$$

(we regard $b^j(X) = 0$ for $j < 0$).

This generalizes as follows:

Consider $U_1, \dots, U_N$ and put

$$U_{(j)} := U_{k_0} \cap \cdots \cap U_{k_j}$$

Proposition 5.1.

$$b^i(U_1 \cup \cdots \cup U_N) \le \sum_{U_{(j)}} b^j(U_{(i-j)})$$

Proof: Note that

$$U_0 \cup \ldots \cup U_{t+1} = (U_1 \cup \cdots U_t) \cup U_{t+1}$$

$$(U_0 \cup \cdots \cup U_t) \cap U_{t+1} = (U_0 \cap U_{t+1}) \cup \ldots \cup (U_t \cap U_{t+1})$$

Apply the previous estimate and induction to the pair $(U_1 \cup \ldots \cup U_t)$, U_{t+1} and use induction to estimate $b^{i-1}((U_0 \cup \ldots \cup U_t) \cap U_{t+1})$.

It is extremely useful to further generalize these estimates to give bounds on the ranks of induced maps on cohomology (note $b^i = \mathrm{rk}(\mathrm{Id}_{H^i})$).

Let $V_1 \xrightarrow{g} V_2 \xrightarrow{f} V_3$ be linear transformations of vector spaces. Then

$$rk(fg) \le \min(rk(f), rk(g)) \ .$$

Thus, if $A\overset{u}{\subset}B\overset{v}{\subset}C\overset{w}{\subset}D$, with u, v, w, the inclusion maps and u^*, v^*, w^*, the induced maps on cohomology, then

$$(*) \qquad rk((uvw)^*) \leq rk(v^*) \ ,$$

Definition 5.2. If $A \subset B$ let $b^i(A, B)$ denote $rk(u^*)$, where $u^* : H^i(B) \to H^i(A)$.

Remark 5.3. If A, B are open, with $\overline{A} \subset B$, then there exists a submanifold, Y^n, with smooth boundary, such that $A \subset Y^n \subset B$. Then

$$b^i(A, B) \leq b^i(Y, Y) = b^i(Y) < \infty \ .$$

Let $\overline{U}_i^j \subset U_i^{j+1}$, $i = 1, \ldots, N$, $j = 0, \ldots, n+1$. Put $X^j = \cup_i U_i^j$. Thus

$$X^0 \subset X^1 \subset \ldots \subset X^{n+1} \ .$$

Then we have the following generalization of Proposition 5.1.

Proposition 5.4.

$$b^i(X^0, X^{n+1}) \leq \sum_{j,(i-j)} b^j(U^j_{(i-j)} U^{j+1}_{(i-j)})$$

The proof of Proposition 5.4, which is a standard application of the *double complex* associated to an open cover, is given in the Appendix to this section.

We are particularly interested in coverings by balls. First we note the obvious

Lemma 5.5.

$$B_r(p_1) \cap \ldots \cap B_r(p_j) \neq \emptyset$$

implies that for $1 \leq i \leq j$

$$B_r(p_1) \cap \ldots \cap B_r(p_j) \subset B_r(p_i) \subset B_{5r}(p_i) \subset B_{10r}(p_1) \cap \ldots \cap B_{10r}(p_j) \ .$$

Proof: This follows immediately from the triangle inequality.

Content.

Put

$$b^i(r, p) := b^i(B_r(p),\ B_{5r}(p))$$

$$\text{cont}(r, p) := \sum_i b^i(r, p)$$

Note. If $r > \text{dia}(M)$, then $b^i(r, p) = b^i(M)$.

Corollary 5.6.

$$b^i(B_r(p_1) \cap \cdots B_r(p_j)\ ,\ B_{10\,r}(p_1) \cap \cdots B_{10\,r}(p_j)) \leq b^i(r, p_i)\ , \quad 1 \leq i \leq j\ .$$

Proof: This follows from Lemma 5.5 and the inequality $(*)$ preceding Definition 5.2.

Assume now that for any ball, $B_r(p)$, we have $B_r(p) \subset \cup_1^N B_\epsilon(p_i)$, with $p_i \in B_r(p)$ and $N \leq N(\epsilon, r)$.

For the next corollary we will need the observation that if $p' \in B_r(p)$, then by the triangle inequality,

$$(+) \qquad B_{10\cdot r/10}(p') = B_r(p') \subset B_{5r}(p) .$$

Corollary 5.7. *If for all $p' \in B_r(p)$ and $j = 1, \ldots, n+1$*

$$\mathrm{cont}(10^{-j} r, p') \leq c ,$$

then

$$\mathrm{cont}(r, p) \leq (n+1) \cdot 2^{N(10^{-(n+1)} r, r)} \cdot c$$

Proof: Take a cover of $B_r(p)$ by balls $B_{10^{-(n+1)} r}(p_i)$ as above. Put $U_i^j = B_{10^{j-(n+1)} r}(p_i)$ and apply Proposition 5.4. The total number of intersections, $U^j_{(i-j)}$, on the right-hand side of the inequality in Proposition 5.4 is at most $(n+1) 2^{N(10^{-(n+1)} r, r)}$. Also, by Corollary 5.6, we certainly have

$$b^j(U^j_{(i-j)}, U^{j+1}_{(i-j)}) \leq c .$$

Since

$$B_r(p) \subset U^0 \subset U^{n+1} \subset B_{5r}(p)$$

(see (+) above) the claim follows from the inequality (∗).

Thus, the content of a given ball can be estimated in terms of the contents of certain smaller balls.

There is an easier, but equally important means of estimating content.

Compression

Definition 5.8. We say $B_r(p)$ compresses to $B_s(q)$ and write $B_r(p) \mapsto B_s(q)$, if

1) $5s + \overline{p, q} \leq 5r$.

2) There is a homotopy, $f_t : B_r(p) \to B_{5r}(p)$, with f_0 the inclusion and $f_1(B_r(p)) \subset B_s(q)$.

Note. By 1), $B_r(p) \mapsto B_s(q)$ implies $s \leq r$.

Lemma 5.9. *If $B_r(p) \mapsto B_s(q)$, then*

$$b^i(r, p) \leq b^i(s, q) .$$

$$\mathrm{cont}(r, p) \leq \mathrm{cont}(s, q) ,$$

Proof: Obvious by (∗).

Rank

Now for each ball, $B_r(p)$ we define (inductively) an integer called the rank. This invariant enables us to conveniently combine our two methods of estimating content (Corollary 5.7 and Lemma 5.9).

Definition 5.10.

i) $\mathrm{rank}(r,p) := 0$, if $B_r(p) \mapsto B_s(q)$, with $B_s(q)$ contractible.

ii) $\mathrm{rank}(r,p) := j$, if $\mathrm{rank}(r,p)$ is *not* $\leq j-1$ and if $B_r(p) \mapsto B_s(q)$, such that for all $q' \in B_s(q)$, with $s' \leq \frac{1}{10}s$, we have $\mathrm{rank}(s',q') \leq j-1$.

Remark 5.11. Of necessity, there exists *some* $B_{s'}(q')$ as in ii) with $\mathrm{rank}(s',q') = j-1$; otherwise we would have $\mathrm{rank}(r,p) \leq j-1$.

Proposition 5.12. *If balls of radius $< \epsilon$ are contractible, then*

$$\mathrm{rank}(r,p) \leq \frac{\log r/\epsilon}{\log 10} + 1 \ .$$

Proof: Trivial by induction.

Corollary 5.13.

$$\mathrm{cont}(r,p) \leq \left((n+1)\,2^{N(10^{-(n+1)}r,r)}\right)^{rank(r,p)} \ .$$

Proof: By induction, this follows from Corollary 5.7, and Definition 5.10.

Remark 5.14. Needless to say, there is some degree of arbitrariness involved in the choice of constants 5, 10, (and $\frac{1}{2}$) which appear in Definitions 5.2, 5.10 (and 6.1) respectively.

Remark 5.15. In certain respects, our terminology and notation in §§5,6, differ somewhat from that of [G].

Appendix. The generalized Mayer-Vietoris estimate.

Our considerations here are very similar to those of [BT], Chapter II.

Let $\cup_i U_i = X$ be an open cover of X. Put

$$C^{i,j} = \oplus_{(i)} C^j(U_{(i)})$$

where $C^j(U_{(i)})$ denotes the space of singular j-cochains, with coefficients in some field. The *double complex* $C^* := \oplus_{i,j} C^{i,j}$ has two differentials,

$$\delta : C^{i,j} \to C^{i+1,j} \ , \quad \delta^2 = 0$$

$$d : C^{i,j} \to C^{i,j+1} \ , \quad d^2 = 0$$

(d is induced by $d : C^j(U_{(i)}) \to C^{j+1}(U_{(i)})$). The *total differential,* $(d+\delta)$, also satisfies $(d+\delta)^2 = 0$.

A j-cochain on $X = \cup U^i$ determines and is determined by $x \in C^{0,j}$, with $\delta x = 0$. Under this identification $d : C^{0,j} \to C^{0,j+1}$ corresponds to $d : C^j(X) \to C^{j+1}(X)$.

A basic fact we need is that C^* is *δ-acyclic for $i > 0$*, i.e. $y \in C^{i,j}$ $(i > 0)$; $\delta y = 0$ implies $y = \delta z$ for some $z \in C^{i-1,j}$.

Let $X^k = \cup_i U_i^k$, where $\overline{U}_i^k \subset U_i^{k+1}$. Denote by $C^*(k)$, the double complex associated to $X^k = \cup_i U_i^k$. There is a natural restriction map $r_k : C^{i,j}(k) \to C^{i,j}(k+1)$, commuting with d, δ.

Proof of Proposition 5.4: Let $Z^j \subset C^{0,j}(j+1) \cap \ker d + \delta$, be a space of representative cocycles, mapping isomorphically onto $H^j(X^{j+1})$. We will define a filtration,

$$Z^j = Z^j_{j+1} \supset Z^j_j \supset \ldots \supset Z^j_0 \ ,$$

such that

$$(\times) \qquad \dim(Z^j_{s+1} \,/\, Z^j_s) \le \sum_{s,(j-s)} b^s(U^s_{(j-s)}, U^{s+1}_{(j-s)}) \ ,$$

and if $z \in Z^j_0$, then $r_0^* \cdot r_1^* \cdots r_j^*(z)$ is *exact*. This will suffice to prove Proposition 5.4.

Put

$$Z^j_j := \{z \in Z^j \mid r_j^*(z) \text{ is } d \text{ exact}\} \ .$$

Choose a linear map $d^{-1}: r_j^*(Z^j_j) \to C^{0,j-1}(j)$, $dd^{-1}(z) = z$. From $\delta z = 0$, $d\delta = -\delta d$, we get $d\delta d^{-1}(z) = 0$ (and of course $\delta(\delta d^{-1}(z)) = 0$).

Define $Z^j_{j-1} \subset Z^j_j$ by

$$Z^j_j := \{z \in Z^j_{j-1} \mid r^*_{j-1}\delta d^{-1} r_j^*(z) \text{ is } d \text{ exact }\}$$

By proceeding in this way, we obtain $Z^j_{j+1} \supset Z^j_j \supset \ldots \supset Z^j_0$, for which the inequality $(\times)$ obviously holds.

Note that $z \in Z^j_0$ implies

$$r_0^* \delta d^{-1} r_1^* \delta d^{-1} \ldots \delta d^{-1} r_j^*(z) \equiv 0 \ ,$$

since an exact 0-cochain vanishes identically.

To show $r_0^* \ldots r_j^*(z)$ is exact, put

$$a_s = (-1)^{s+j-1} r_0^* \ldots r_{s-1}^* (d^{-1} r_s^*)(\delta d^{-1} r_{s+1}^* \ldots \delta d^{-1} r_j^*)(z) \ .$$

Then

$$r_0^* \ldots r_j^*(z) = (d+\delta)(a_{j-1} + \ldots + a_0) \ .$$

Using δ-acyclicity, choose $b_0 \in C^{j-1,0}(0)$, with $\delta b_0 = a_0$. Put $a_1' = a_1 - db_0$. Then

$$\begin{aligned} r_0^* \ldots r_j^*(z) &= (d+\delta)(a_{j-1} + \ldots + a_0 - (d+\delta)b_0) \\ &= (d+\delta)(a_{j-1} + \ldots + a_2 + a_1') \end{aligned}$$

Proceeding in this way, we find by induction, $\hat{a}_{j-1} \in C^{0,j-1}(0)$, with

$$r_0^* \cdot \ldots r_j^*(z) = (d+\delta)\hat{a}_{j-1} \ .$$

Then, we have

$$r_0^* \cdots r_j^*(z) = d\hat{a}_{j-1} \ ,$$
$$0 = \delta\hat{a}_{j-1} \ ,$$

which completes the proof.

6. Rank, curvature and diameter.

In this section, we will show how a lower bound on curvature leads to an estimate on rank(r,p), and hence via Corollary 5.13, to an estimate on cont(r,p). For this purpose, it is convenient to work with a slightly modified definition of rank.

Definition 6.1. A ball, $B_r(p)$, is called *incompressible* if $B_r(p) \mapsto B_s(q)$ implies $s > \frac{1}{2}r$.

It is obvious that any ball, $B_r(p)$, can be compressed either to a contractible ball (in which case rank$(r,p) = 0$) or to an incompressible ball.

Definition 6.2.

i) rank$'(p,r) := 0$, if $B_r(p) \mapsto B_s(q)$, with $B_s(q)$ contractible

ii) rank$'(p,r) := j$ if rank$'(p,r) \neq j-1$ and $B_r(p) \mapsto B_s(q)$ such that: $B_s(q)$ is incompressible and for all $q' \in B_s(q)$ and $s' \leq \frac{1}{10}s$, we have rank$(q',s') \leq j-1$.

Thus, we have modified Definition 5.10, by adding the stipulation that the ball, $B_s(q)$, of ii), must be *incompressible.*

Clearly, rank$'(p,r)$ still satisfies the bound of Proposition 5.12. Moreover, it is obvious that

$$\text{rank}(r,p) \leq \text{rank}'(r,p) \ .$$

The reason for insisting on incompressibility in the definition of rank$'(r,p)$ stems from

Lemma 6.3. *Let $B_r(p) \subset M^n$, a complete riemannian manifold. Assume*

$$5s + \overline{p,y} \leq 5r \ ,$$
$$\overline{p,y} \leq 2r \ .$$

Then if $B_r(p)$ does not compress to $B_s(y)$, there exists a critical point, x, of y, with

$$s \leq \overline{x,y} \leq r + \overline{p,y} \ .$$

Thus, $x \subset B_{r+2\overline{p,y}}(p) \subset B_{5r}(p)$.

Proof: If there were no such critical point, then by the Isotopy Lemma 1.4, the ball $B_{r+\overline{p,y}}(y)$ could be deformed to lie inside of $B_s(y)$. Since, $5s + \overline{p,y} \leq 5r$, and

$$B_r(p) \subset B_{r+\overline{p,y}}(y) \subset B_{5r}(p) \ ,$$

this would contradict the assumption that $B_r(p)$ does not compress to $B_s(y)$.

Now we can show a connection between the size of $\mathrm{rank}'(r,p)$ and the existence of critical points.

We will need the observation that if $p' \in B_r(p)$, then by the triangle inequality,

$$(*) \qquad B_{5\cdot(r/10)}(p') = B_{r/2}(p') \subset B_{3r/2}(p) \ .$$

Lemma 6.4. *Let M^n be riemannian and let $\mathrm{rank}'(r,p) = j$. Then there exists $y \in B_{5r}(p)$ and $x_j, \ldots, x_1 \in B_{5r}(p)$, such that for all $i \le j$, x_i is critical with respect to y and*

$$\overline{x_i, y} \ge \frac{5}{4}\overline{x_{i-1}, y} \ .$$

Proof: We can assume without loss of generality that $B_r(p)$ is incompressible (in this case, we will see that $y \in B_{3r/2}(p)$, $x_i \in B_{3r/2}(p)$). Put $p_j = p$, $r_j = r$. By the definition of $\mathrm{rank}'(r,p)$, there exists $\hat{p}_{j-1} \in B_{r_j}(p_j)$, $\hat{r}_{j-1} \le \frac{1}{10} r_j$, such that

$$\mathrm{rank}'(\hat{r}_{j-1}, \hat{p}_{j-1}) = j - 1 \ .$$

By $(*)$ above,

$$B_{3r_j/2}(p_j) \supset B_{5r_{j-1}}(\hat{p}_{j-1}) \ .$$

If $B_{\hat{r}_{j-1}}(\hat{p}_{j-1})$ is incompressible, put

$$p_{j-1} = \hat{p}_{j-1} \ , \quad r_{j-1} = \hat{r}_{j-1} \ .$$

If not, there exists an incompressible ball, which in this case we call $B_{r_{j-1}}(p_{j-1})$, such that $B_{\hat{r}_{j-1}}(\hat{p}_{j-1}) \mapsto B_{r_{j-1}}(p_{j-1})$ and

$$\mathrm{rank}'(r_{j-1} \ , p_{j-1}) = j - 1 \ .$$

Since $B_{\hat{r}_{j-1}}(\hat{p}_{j-1}) \mapsto B_{r_{j-1}}(p_{j-1})$ implies $B_{5r_{j-1}}(p_{j-1}) \subset B_{5\hat{r}_{j-1}}(\hat{p}_{j-1})$, in either case we obtain

$$B_{3r_j/2}(p_j) \supset B_{5r_{j-1}}(p_{j-1})$$

Also, since in the second case $r_{j-1} \le \hat{r}_{j-1}$, in either case, we have

$$r_{j-1} \le \frac{1}{10} r_j \ .$$

By proceeding in this fashion, we obtain balls, $B_{r_i}(p_i)$, $i = 0, 1, \ldots, j$, such that for $1 \le i \le j$, $B_{r_i}(p_i)$ is incompressible and

$$B_{3r_i/2}(p_i) \supset B_{5r_{i-1}}(p_{i-1})$$
$$r_{i-1} \le \frac{1}{10} r_i \ .$$

Put $y = p_0$. Then, $y \in B_{3r_i/2}(p_i)$, for all $1 \le i \le j$. In particular,

$$\overline{p_i, y} + 5 \cdot \frac{1}{2} r_i \le 4r_i < 5r_i \ ,$$
$$\overline{p_i, y} \le \frac{3}{2} r_i < 2r_i$$

(the conditions of Lemma 6.3).

Since, $B_{r_i}(p_i)$ is incompressible, it does not compress to $B_{r_i/2}(y)$. Thus, by Lemma 6.3, there exists a critical point, x_i, with

$$\frac{1}{2} r_i \le \overline{x_i, y} \le r_i + 2 \cdot \frac{3}{2} r_i = 4r_i \ .$$

Then,

$$\begin{aligned} \overline{x_i, y} &\ge \frac{1}{2} r_i \\ &\ge 5r_{i-1} \\ &\ge \frac{5}{4} 4r_{i-1} \\ &\ge \frac{5}{4} \overline{x_{i-1}, y} \ . \quad \text{q.e.d.} \end{aligned}$$

Corollary 6.5.

$$\operatorname{rank}(r,p) \le \begin{cases} \mathcal{N}(n) & H = 0 \\ \mathcal{N}(n, Hd^2) & H < 0 \ . \end{cases}$$

Proof: This follows immediately from Corollary 2.7, Lemma 6.4 and the inequality $\operatorname{rank}(r,p) \le \operatorname{rank}'(r,p)$.

Proof of Theorem 3.8: By Proposition 3.11,

$$N(10^{-(n+1)} r, r) \le \begin{cases} N_1(n,\ 10^{-(n+1)}) \\ N_1(n,\ Hd^2,\ 10^{-(n+1)}) \end{cases}$$

with $N_1(10^{-(n+1)} r, r)$ the covering number appearing in Corollary 5.7. Hence, by that corollary, and by Corollary 6.5 for all $\epsilon > 0$,

$$\begin{aligned} \sum b^i(M^n) &= \operatorname{cont}(d+\epsilon, p) \\ &\le (\,(n+1)\, 2^{N_1})^{\mathcal{N}} \end{aligned}$$

q.e.d.

Remark 6.6. Inspection of the bounds given in Corollary 2.7 and Proposition 3.11 (compare Corollary 2.11) reveals that the dependence on n of the constant $C(n)$ in Theorem 3.8 is at worst of the form $2^{2^{an}}$ (for suitable $a > 0$). However, Abresch has shown that by arguing more carefully (along essentially the same lines as we have done) one obtains $C(n) \le 2^{an^3}$; [A], [Me]. Recall that in view of the existence of flat tori, $C(n) = 2^n$ is the best one could hope for.

7. Ricci curvature, volume and the Laplacian.

In this section we present some basic properties of manifolds whose Ricci curvature satisfies $\mathrm{Ric}_{M^n} \geq (n-1)H$. In particular, after proving Proposition 3.11, we derive some estimates involving the Laplacian, which are used in §8. There, we prove a theorem of Abresch-Gromoll, asserting that complete manifolds with $\mathrm{Ric}_{M^n} \geq 0$, which satisfy certain additional conditions, have finite topological type.

Let $\{e_i\}$ be an orthonormal basis of M^n_p. We denote by $\mathrm{Ric}(u,v)$ the symmetric bilinear form,

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

Thus, $\mathrm{Ric}(u,v)$ is the trace of the linear transformation $w \to R(w,u)v$.

We write

$$\mathrm{Ric}_{M^n} \geq (n-1)H \ ,$$

if

$$\mathrm{Ric}(v,v) \geq (n-1)H \ ,$$

for all unit tangent vectors v. Of course, this condition is implied by $K_{M^n} \geq H$, but not vice versa.

Suppose $\gamma|[0,\ell]$ contains no cut point. Then the distance function, $r = \rho_{\gamma(0)}$, is smooth near $\gamma\,|\,[0,\ell]$. Put $N = \mathrm{grad}\, r$. Thus, $N(\gamma(t)) = \gamma'(t)$. Let $e_2, \cdots e_n$ be orthonormal, with

$$\langle e_i, \gamma'(0)\rangle_{\gamma(0)} = 0 \ ,$$

$$\nabla_N e_i = 0 \ .$$

Then

$$\mathrm{Ric}(N,N) = \sum_i \langle (\nabla_{e_i}\nabla_N - \nabla_N\nabla_{e_i} - \nabla_{[e_i,N]})N, e_i\rangle \ .$$

We have,

$$\nabla_N N \equiv 0 \ .$$

Also,

$$\begin{aligned} -\sum_i \langle \nabla_N \nabla_{e_i} N, e_i\rangle &= -\sum_i N\langle \nabla_{e_i} N, e_i\rangle \\ &= -m' \ , \end{aligned}$$

where m is the mean curvature of the distance sphere, $\partial B_r(\gamma(0))$, in the direction of the inner normal, $-N$ (and $m' = \dfrac{\partial m}{\partial r}$). Finally,

$$\begin{aligned} -\sum \langle \nabla_{[e_i,N]} N, e_i\rangle &= -\sum_{i,j} \langle \nabla_{e_i} N, e_j\rangle \ \langle \nabla_{e_j} N, e_i\rangle \ , \\ &= -\|\mathrm{Hess}_r\|^2 \ , \end{aligned}$$

where Hess_r denotes the Hessian of r. Thus, we get the basic equation,

$$(*) \qquad \|\mathrm{Hess}_r\|^2 + \mathrm{Ric}(N,N) = -m' \ .$$

Additionally, letting Δ denote the Laplacian, we have

$$\begin{aligned}\Delta r &= \Big(\sum_2^n e_i e_i + NN - \nabla_{e_i} e_i - \nabla_N N\Big) r \\ &= m\ .\end{aligned}$$

Alternatively, this relation follows from the formula, for Δ in geodesic polar coordinates,

$$\Delta = \frac{\partial^2}{\partial r^2} + m(x)\frac{\partial}{\partial r} + \widetilde{\Delta}\ ,$$

where $\widetilde{\Delta}$ is the intrinsic Laplacian of the distance sphere $\partial B_r(\gamma(\sigma)\,)$.

An invariant definiton of the Laplacian is,

$$\Delta f = \operatorname{tr}(\mathrm{Hess}_f)\ .$$

Thus, we also have

$$\operatorname{tr}(\mathrm{Hess}_r) = m\ .$$

From the Schwarz inequality and the fact that one eigenvalue of Hess_r is $\equiv 0$ (corresponding to the eigenvector, N), we get

$$(**) \qquad \|\mathrm{Hess}_r\|^2 \geq \frac{m^2}{n-1}\ .$$

Substituting $(*)$ into $(**)$ gives the differential inequality,

$$\mathrm{Ric}(N,N) \leq -\frac{m^2}{n-1} - m'\ .$$

Note that, as $r \to 0$,

$$m(r) \sim \frac{n-1}{r}\ .$$

Set $u = \dfrac{n-1}{m}$. Then if we assume $\mathrm{Ric}_{M^n} \geq (n-1)H$, we easily obtain

$$\frac{u'}{1+Hu^2} \geq 1\ .$$

By integrating this expression, we find that $\mathrm{Ric}_{M^n} \geq (n-1)H$ implies

$$(+) \qquad m(x) \leq m_H(r(x)\,)\ ,$$

or equivalently,

$$(++) \qquad \Delta r(x) \leq \Delta_H r\big|_{r=r(x)}\ ,$$

where $m_H(r)$, the mean curvature of $\partial B_r(\underline{p}) \subset M_H^n$ in direction $-N$, is given by

$$m_H(r) = (n-1)\begin{cases} \sqrt{H}\cot\sqrt{H}\,r & H>0 \\ r^{-1} & H=0 \\ \sqrt{-H}\coth\sqrt{-H}\,r & H<0 \end{cases}$$

Let ω denote the volume form on the unit sphere. Write

$$dr \wedge \mathcal{A}(r)\omega$$

for the volume form on $M^n \setminus C_p$, Then $m = \frac{\mathcal{A}'}{\mathcal{A}}$. Differentiating $\mathcal{A}/\mathcal{A}_{H,n-1}$ and using $(+)$ gives

$$[\mathcal{A}(r)/\mathcal{A}_{H,n-1}(r)] \downarrow ,$$

$$\mathcal{A}(r) \leq \mathcal{A}_{H,n-1}(r) .$$

($\mathcal{A}_{H,n-1}(r)$ is defined prior to Lemma 4.3.)

Now, by arguing as in the proofs, of Lemma 4.3 and Proposition 4.7, we can immediately extend Lemma 4.3 as follows.

Proposition 7.1. *Let* $\mathrm{Ric}_{M^n} \geq (n-1)H$ *and let* $X \subset M^n$ *be compact. Then for* $r_1 < r_2$,

$$\frac{\mathrm{Vol}(T_{r_1}(X))}{\mathrm{Vol}(T_{r_2}(X))} \geq \frac{V_{n,H}(r_1)}{V_{n,H}(r_2)} .$$

Remark 7.2. In the basic case, $X = p$, the above inequality was emphasized in [G]; compare also [C2].

Proof of Proposition 3.11: Take a maximal set of points, p_i, in $\overline{B_{r-\epsilon/2}(p)}$ at mutual distance $\geq \frac{\epsilon}{2}$. Clearly $\{p_i\}$ is $\frac{\epsilon}{2}$-dense in $\overline{B_{r-\epsilon/2}(p)}$, and hence, ϵ-dense in $B_r(p)$. The balls $\{B_{\epsilon/4}(p_i)\}$ are all disjoint. Moreover, by Proposition 7.1,

$$\frac{V_{n,H}(\epsilon/4)}{V_{n,H}(2r)} \leq \frac{\mathrm{Vol}(B_{\epsilon/4}(p_i))}{\mathrm{Vol}(B_{2r}(p_i))} ,$$

while since $B_r(p) \subset B_{2r}(p_i)$,

$$\frac{\mathrm{Vol}(B_{\epsilon/4}(p_i))}{\mathrm{Vol}(B_{2r}(p_i))} \leq \frac{\mathrm{Vol}(B_{\epsilon/4}(p_i))}{\mathrm{Vol}(B_r(p))} .$$

Thus, the number of balls is bounded by

$$\frac{V_{n,H}(2r)}{V_{n,H}(\epsilon/4)} .$$

If $B_\epsilon(p_j) \cap B_\epsilon(p_i) \neq \emptyset$, then $B_\epsilon(p_j) \subset B_{3\epsilon}(p_i)$. Then, as above, it follows that the multiplicity of our covering is bounded by

$$\frac{V_{n,H}(3\epsilon)}{V_{n,H}(\epsilon/2)} . \qquad \text{q.e.d.}$$

Now, the scale invariant inequalities of Proposition 3.11 follow by an obvious scaling argument.

Remark 7.3. At this point it is clear that the hypothesis, $\mathrm{Ric}_{M^n} \geq (n-1)H$, implies that the map in Proposition 4.7 is volume decreasing.

We now observe that the inequality, $(++)$, on the Laplacian of the distance function, can be generalized in a meaningful way so as to include points which lie on the cut locus. This discussion goes back to a fundamental paper of E. Calabi, [Ca].

First we need some definitions.

Definition 7.4. An *upper barrier* for a continuous function f at the point x_0, is a C^2 function, g, defined in some neighborhood of x_0, such that $g \geq f$ and $g(x_0) = f(x_0)$.

The crucial observation for applications to geometry is the following.

Lemma 7.5. *If $\gamma(\ell)$ is a cut point of $\gamma(0)$, then for all $\epsilon < \ell$, $\rho_{\gamma(\epsilon)}(x) + \epsilon$ is an upper barrier for $r = \rho_{\gamma(0)}$ at $\gamma(\ell)$.*

Proof: This follows immediately from the triangle inequality.

Definition 7.6. We say $\Delta f(x_0) \leq a$ (and $\Delta(-f(x_0)) \geq -a$) in the *barrier sense* if for all $\epsilon > 0$, there is an upper barrier $f_{x_0,\epsilon}$ for f at x_0 with

$$\Delta f_{x_0,\epsilon} \leq a + \epsilon \ .$$

Now we can generalize (++) above as follows.

Proposition 7.7. *Let M^n be complete, with*

$$\mathrm{Ric}_{M^n} \geq (n-1)H \ .$$

i) *If $f(r)$ satisfies, $f' \geq 0$, then in the barrier sense,*

$$\Delta f(r(x)) \leq \Delta_H f(r)\big|_{r=r(x)} \ .$$

ii) *If $f(r)$ satisfies $f' \leq 0$, then in the barrier sense,*

$$\Delta f(r(x)) \geq \Delta_H f(r)\big|_{r=r(x)} \ .$$

Proof: It suffices to prove i). At smooth (i.e. noncut) points, it is clear from (+) and the formula for Δ in polar coordinates. At cut points, it follows immediately by using the barrier $f(\rho_{\gamma(\epsilon)}(x) + \epsilon)$.

Functions which satisfy say $\Delta f \geq 0$ in the barrier sense, also satisfy a maximum principle. This fact (due to Calabi) was used by Eschenberg and Heintze [EH] to give a very short proof of the splitting theorem of [CGl1]. (They also gave a somewhat longer, but completely elementary proof along closely related lines). Theorem 7.9 below, which is crucial for the discussion of §8, was partly inspired by their work.

Theorem 7.8 (Maximum principle). *Let M be a connected riemannian manifold and let $f \in C^0(M)$. Suppose that $\Delta f \geq 0$ in the barrier sense. Then f attains no weak local maximum unless it is a constant function.*

For completeness, in the Appendix to this section, we give a proof of Theorem 7.8.

We now give an estimate of Abresch-Gromoll on the growth of nonnegative Lipschitz functions whose Laplacian is bounded above in the barrier sense. This will be applied to excess functions, i.e. functions of the form

$$e(x) = \overline{x,p} + \overline{x,q} - \overline{p,q}$$

(for fixed p, q). The estimate involves a comparison function on the model space M^n_H. We now specify this function.

Given $\underline{y} \in M^n_H$, as usual, put $r(\underline{x}) := \rho_{\underline{y}}(\underline{x})$. Fix $R > 0$ and a constant $b > 0$. Then there is a unique smallest function, $G(r(\underline{x}))$, defined on $M^n_H \setminus \underline{y}$ satisying,

1) $G > 0$ $(0 < r < R)$

2) $G' < 0$ $(0 < r < R)$

3) $G(R) = 0$,

4) $\Delta_H G \equiv b$,

For any H, the function G can be written in closed form; see [AGl]. Here we need only the case, $H = 0$, $n > 2$. Then,

$$G(r) = \frac{b}{2n}(r^2 + \frac{2}{n-2}R^n r^{2-n} - \frac{n}{n-2}R^2)\ .$$

Let M^n be complete with $\mathrm{Ric}_{M^n} \geq 0$, $y \in M^n$, $r(x) = \rho_y(x)$. Then by 2) and 4) together with Proposition 7.7,

$$(\times) \qquad\qquad \Delta G(r(x)) \geq b$$

holds in the sense of barriers.

For f a Lipschitz function on M^n, denote by dil f the smallest constant k such that for all x_1, x_2,

$$|f(x_1) - f(x_2)| \leq k\,\overline{x_1, x_2}\ .$$

Theorem 7.9. *Let M^n be complete, with* $\mathrm{Ric}_{M^n} \geq (n-1)H$. *Let $u : B_{R+\eta}(y) \to R$ (for some $\eta > 0$) be a Lipschitz function satisfying*

i) $u \geq 0$,

ii) $u(y_0) = 0$,

for some $y_0 \in \overline{B_R(y)}$.

iii) dil $u \leq a$,

iv) $\Delta u \leq b$,

in the barrier sense. Then for all c, with $0 < c < R$

$$u(y) \leq a \cdot c + G(c)\ .$$

Proof: Take $\epsilon < \eta$ and define the function G using the value $R + \epsilon$, in place of R. Since we can eventually let $\epsilon > 0$, it will suffice to prove the inequality in this case.

In what follows we write G for $G(r(x))$. Fix $0 < c < R$ and suppose the bound is false. Then by iii), it follows that

$$u \mid \partial B_c(y) \geq G \mid \partial B_c(y)$$

Also, by i) and property 3) of the function G,

$$u \mid \partial B_{R+\epsilon}(y) \geq G \mid \partial B_{R+\epsilon}(y) \; .$$

Thus, the function $(G-u)$ satisfies

$$(G-u) \mid \partial B_c(y) \leq 0 \; ,$$

$$(G-u) \mid \partial B_{R+\epsilon}(y) \leq 0 \; .$$

However, by ii) and property 1) of G,

$$(G-u)(y_0) > 0 \; .$$

Hence, $(G-u) \mid \overline{B_{R+\epsilon}(y)} \setminus B_c(y)$ has a strict interior maximum. But since by iv) and $(\times)$ we know

$$\Delta(G-u) \geq 0$$

holds in the barrier sense, this contradicts the maximum principle (Theorem 7.8).

Remark 7.10. One easily checks that in the explicit formula for G in the case $H=0$, the optimal value of c is the unique number satisfying $0 < c < R$, and

$$c(\,(R/c)^n - 1\,) = \frac{an}{b} \; .$$

However, in Corollary 7.11 below, a value which is approximately optimal is all that is required.

As previously mentioned, Theorem 7.9 can be used to obtain an estimate on excess functions. Let

$$E(x) = \overline{x, y_1} + \overline{x, y_2} - \overline{y_1, y_2}$$

be the excess function associated to $y_1, y_2 \in M^n$. We can regard $E(x)$ as the excess of any triangle with vertices x, y_1, y_2 and all sides minimal.

Let γ be a minimal geodesic from y_1 to y_2. The function E satisfies

i) $E(x) \geq 0$,

ii) $E \mid \gamma \equiv 0$.

iii) dil $E \leq 2$.

In case $\mathrm{Ric}_{M^n} \geq 0$, by (++) above and Proposition 7.7, we have

iv) $\Delta E \leq (n-1)(\dfrac{1}{s_1} + \dfrac{1}{s_2})$,
where $s_j(x) = \overline{x, y_j}$.

Define the function $s(x)$ by

$$s(x) = \min(s_1(x), s_2(x)\,) \; .$$

Define the height function, $h(x)$, by

$$h(x) = \min_{\gamma, t} \overline{x, \gamma(t)} \; ,$$

where γ is a minimal segment from y_1 to y_2.

Clearly, $h(x) = 0$ implies $E(x) = 0$ and, by the triangle inequality,

$$E(x) \leq 2h(x) \ .$$

Under the assumption $\mathrm{Ric}_{M^n} \geq 0$, we now derive a quantitative relation between the values $h(x)$ and $s(x)$ which guarantees that $E(x)$ is small at the point x.

Corollary 7.11. *If* $\mathrm{Ric}_{M^n} \geq 0$, *then for* $h \leq \frac{1}{2}s$,

$$\begin{aligned} E &\leq 8\left(\frac{h^n}{s}\right)^{1/(n-1)} \\ &= 8\left(\frac{h}{s}\right)^{1/(n-1)} h \ . \end{aligned}$$

Proof: The function E satisfies the hypothesis of Theorem 7.9 with $a = 2$, $b = 4(n-1)/s$. Use $c = (2h^n/s)^{1/(n-1)}$ in the estimate

$$E \leq 2c + G(c) \ .$$

The sum of the first and third terms in the explicit expression for G is negative. The middle term gives a contribution at most equal to

$$\frac{2(n-1)}{sn} \cdot \frac{2}{n-2} h^n \left(\frac{2h^n}{s}\right)^{(2-n)/(n-1)} \leq 4c \ .$$

The claim follows immediately.

Remark 7.12. The estimate we have derived is of particular interest at points x, where h, s are large individually, but h^n/s is small. Roughly speaking, such triangles might be called "thin".

Remark 7.13. Suppose, in fact, that $K_{M^n} \geq 0$. Let y_0 be a closest point to x among all points which lie on minimal segments from y_1 to y_2. Divide a triangle with vertices x, y_1, y_2 into two right triangles with vertices x, y_1, y_0 and x, y_2, y_0. Put $\overline{y_0, y_j} = t_j$, $j = 1, 2$. Then Toponogov's theorem B) gives

$$(\surd) \qquad \begin{aligned} s_j &\leq (h^2 + t_j^2)^{1/2} \ , \\ &\leq t_j\left(1 + 2\left(\frac{h}{t_j}\right)^2\right) \ . \end{aligned}$$

Thus, for $t = \min(t_1, t_2)$, we have

$$E \leq \left(\frac{h}{t}\right) \cdot h \ .$$

If, at the point x, the values h, t are large but h^2/t is small, then E is still small. Note that such *thin* triangles are *not* required to be *as thin* as those in Remark 7.12.

Remark 7.14. Corollary 7.11 is the first estimate in which a nontrivial bound on a *sum of distances*, $s_1 + s_2$, is obtained from a bound on Ricci curvature. But at present, there is no useful bound on the individual s_j, as in $(\surd)$.

Appendix. The maximum principle.

Proof of Theorem 7.8: Let p be a weak local maximum i.e. $f(p) \geq f(x)$ for all x near p. Take a small normal coordinate ball, $B_\delta(p)$, and assume that there exists $z \in \partial B_\delta(p)$ such that $f(p) > f(z)$. Then, by continuity, $f(p) > f(z')$ for $z' \in \partial B_\delta(p)$ sufficiently close to z. Choose a normal coordinate system, $\{x_i\}$, such that $z = (\delta, 0, \cdots, 0)$. Put

$$\phi(x) = x_1 - d(x_2^2 + \cdots + x_n^2) ,$$

where d is so large that if $y \in \partial B_\delta(p)$ and $f(y) = f(p)$, then $\phi(y) < 0$. Note that grad ϕ doesn't vanish.

Put

$$\psi = e^{a\phi} - 1 .$$

Then

$$\Delta(\psi) = (a^2 \|\text{grad } \phi\|^2 + a\,\Delta\phi)\,e^{a\phi} .$$

Thus, for a sufficiently large,

$$\Delta\psi > 0 .$$

Moreover,

$$\psi(p) = 0 .$$

For $\eta > 0$, sufficiently small,

$$(f + \eta\psi)|\partial B_\delta(p) < f(p) .$$

Thus, $f + \eta\psi$ has an interior maximum at some point $q \in B_\delta(p)$.

If $f_{q,\epsilon}$ is a barrier for f at q with $\Delta f_{q,\epsilon} \geq -\epsilon$, then $f_{q,\epsilon} + \eta\psi$ is also a barrier for $f + \eta\psi$ at q. For ϵ sufficiently small, we have

$$\Delta(f_{q,\epsilon} + \eta\psi) > 0 .$$

Since $f + \eta\psi$ has a local maximum at q, and

$$f_{q,\epsilon} + \eta\psi \leq f + \eta\psi ,$$

$$(f_{q,\epsilon} + \eta\psi)(q) = (f + \eta\psi)(q) ,$$

we find that $f_{q,\epsilon} + \eta\psi$ has a local maximum at q as well. But this is incompatible with $\Delta(f_{q,\epsilon} + \eta\psi) > 0$. (Note that in normal coordinates at q, $\Delta = \partial_1^2 + \cdots + \partial_n^2$).

It follows that for all small δ, we have $f \mid \partial B_\delta(p) \equiv f(p)$. Since M is connected this implies $f \equiv f(p)$.

8. Nonnegative Ricci curvature, diameter growth and finiteness of topological type

In this section we prove that if M^n is complete, $\mathrm{Ric}_{M^n} \geq 0$, $K_{M^n} \geq -1$, and if a certain additional condition holds, then has finite topological type i.e. M^n is homeomorphic to the interior of a compact manifold with boundary.

The most general form of the additional condition uses the *ray density function*, $\mathcal{R}(r,p)$, associated to $p \in M^n$. However in some ways, a stronger condition formulated in terms of a second function, $\mathcal{D}(r,p)$, the *diameter growth function* is more natural.

Definition 8.1. Let M^n be complete. A *ray* is a geodesic, $\gamma : [0,\infty) \to M^n$, each segment of which is minimal.

When M^n is complete and noncompact then rays always exist.

Proposition 8.2. *Let M^n be complete noncompact. Then for all p, there exists at least one ray, γ, with $\gamma(0) = p$.*

Proof: Since M^n is not compact, there is a sequence, q_i, with $\overline{p,q_i} \to \infty$. Let γ_i be minimal from p to q_i and let $\{\gamma_j\}$ be a subsequence such that $\gamma_j'(0) \to v$, for some $v \in M_p^n$, with $\|v\| = 1$. Let $\gamma : [0,\infty) \to M^n$ be the geodesic with $\gamma'(0) = v$. Then each segment, $\gamma \mid [0,\ell]$, is a limit of minimal segments, $\gamma_j \mid [0,\ell]$, and hence is minimal itself. Thus γ is a ray.

Let $x \in M^n$, and let γ be a ray from p. Note that if

$$t_1 \leq \overline{x,p} - \overline{x,\gamma}$$
$$t_2 \geq \overline{x,p} + \overline{x,\gamma}$$

then

$$\overline{x,\gamma} = h(x) \ ,$$

where $h(x)$ is the height function of §7, for the excess function associated to the points $\gamma(t_1)$, $\gamma(t_2)$.

Definition 8.3. Define the ray density function by

$$\mathcal{R}(r,p) = \sup_{x \in \partial B_r(p)} \{\inf_{\gamma} \overline{x,\gamma} \mid \gamma \text{ a ray}, \gamma(0) = p\} \ .$$

Proposition 8.4. *Let M^n be complete, with $\mathrm{Ric}_{M^n} \geq 0$ on $M^n \setminus B_\lambda(p)$, for some $\lambda < \infty$. Assume*

$$\mathcal{R}(r,p) = o(r^{1/n}) \ .$$

Then for all $\epsilon > 0$ there exists $\delta > 0$, such that if x, γ are as in Definition 8.3 (x arbitrary) with

$$\overline{x,p} \geq \delta^{-1} \ ,$$

and t is sufficiently large relative to $\overline{x,p}$, then the excess function associated to $\gamma(0)$, $\gamma(t)$ satisfies

$$E(x) \leq \epsilon \ .$$

Proof: For the case in which $\mathrm{Ric}_{M^n} \geq 0$ on all of M^n, this is immediate from Corollary 7.11. The general case is dealt with in the Appendix to this section.

On the other hand, in case $K_M \geq -1$, the following proposition provides a positive lower bound for $E(x)$ when x is critical with respect to say y_1.

Put $\overline{x, y_1} = x$.

Proposition 8.5. *Let M^n be complete with $K_{M^n} \geq -1$ $(H < 0)$. Let x be critical with respect to y_1. Then for all $\epsilon > 0$ there exists δ such that*

$$\overline{x, y_2} \geq \delta^{-1}$$

implies

$$E(x) \geq \ell n\left(\frac{2}{1+e^{-2s}}\right) - \epsilon \ .$$

Proof: By Toponogov's theorem B) and the assumption that x is critical with respect to y_1, it suffices to assume that $x, y_1, y_2 \in M^2_{-1} \subset M^n_{-1}$ and that the minimal geodesics from x to y_1, y_2 make an angle $\pi/2$. By hyperbolic trigonometry

$$\cosh \overline{y_1, y_2} = \cosh s \cosh \overline{x, y_2}$$

By the triangle inequality,

$$|\overline{y_1, y_2} - \overline{x, y_2}| \leq s \ .$$

As both $\overline{y_1, y_2}$, and $\overline{x, y_2} \to \infty$ with s fixed,

$$\frac{\cosh \overline{x, y_1}}{\cosh \overline{y_1, y_2}} \to e^{\overline{x, y_1} - \overline{y_1, y_2}}$$

The claim follows easily.

By combining Propositions 8.4 and 8.5 we obtain

Theorem 8.6 (Abresch-Gromoll). *Let M^n be complete with*

i) $\mathrm{Ric}_{M^n} \geq 0$ *on* $M^n \setminus B_\lambda(p)$, *for some* λ,

ii) $\mathcal{R}(r, p) = o(r^{1/n})$, *for some* $p \in M^n$,

iii) $K_{M^n} \geq H > -\infty$.

Then there exists a compact set, C, such that $M^n \setminus C$ contains no critical points of p. In particular, M^n has finite topological type.

We now define the diameter growth functin $\mathcal{D}(r, p)$. For every r, the open set $M^n \setminus \overline{B_r(p)}$ contains only *finitely many unbounded components*, U_r. Each U_r has *finitely many* boundary components, $\Sigma_r \subset \partial B_r(p)$. In particular Σ_r is a closed subset.

Let $\mathrm{dia}(\Sigma_r)$ denote maximum distance, *measured in M^n*, between a pair of points of Σ_r.

Definition 8.7.

$$\mathcal{D}(r, p) = \sup_{\Sigma_r} \mathrm{dia}(\Sigma_r) \ .$$

Given any boundary component, Σ_r, we can construct a ray, γ, such that $\gamma(t) \subset U_r$, for $t > r$ and so, $\gamma(r) \in U_r$. To do so, it suffices to choose the sequence of points, $\{q_i\}$, of Proposition 8.3 to lie in U_r. Then the convergent subsequence γ_j satisfies the conditions above. Hence γ satisfies them as well.

With this observation, it follows immediately from the proof of Theorem 8.6, that if we assume

$$\mathcal{D}(r,p) = o(r^{1/n}) \ ,$$

then for $r \geq r_0$ sufficiently large, no point of any set Σ_r is critical with respect to p.

Fix r_0, U_{r_0}, a boundary component Σ_{r_0} and a ray γ, with $\gamma(r_0) \in \Sigma_{r_0}$, $\gamma(t) \in U_{r_0}$, for $t > r_0$. For each $t \geq r_0$, let Σ_t denote the boundary component of the unbounded component of $M^n \setminus \overline{B_t(p)}$ with $\gamma(t) \in \Sigma_t$. Using the observation of the previous paragraph and the Isotopy Lemma 1.4, we easily construct an imbedding,

$$\psi : (r_0, \infty) \times \Sigma_0 \to U_x \ ,$$

such that

$$\psi(\,(t, \Sigma_{r_0})\,) = \Sigma_t \ .$$

It follows easily that $\psi((r_0,\infty) \times \Sigma_{r_0}))$ is open and closed in U_{r_0}. Hence

$$\psi(\,(r_0,\infty) \times \Sigma_{r_0}) = U_{r_0} \ .$$

Thus we obtain

Theorem 8.8 (Abresch-Gromoll). *Let M^n be complete with*

i) $\mathrm{Ric}_{M^n} \geq 0$ *on* $M^n \setminus B_\lambda(p)$, *for some* λ,

ii) $\mathcal{D}(r,p) = o(r^{1/n})$ *for some* $p \in M^n$,

iii) $K_{M^n} \geq -H > -\infty$.

Then

$$\mathcal{R}(r,p) = o(r^{1/n}) \ .$$

Thus there exists a compact set C such that $M^n \setminus C$ contains no critical points of p. In particular, M^n has finite topological type.

Remark 8.9. Clearly, any two points on $\partial B_r(p)$ can be joined by a broken geodesic passing through p of length $2r$. Thus, one always has

$$\mathcal{D}(r,p) \leq 2r \ ,$$

for the function, $\mathcal{D}(r,p)$, defined in Definition 8.7. However, it is also of interest to consider modified definitions of diameter growth, for which the above inequality, need not hold. For such definitions and their geometric significance, see [AGl], [Liu], [Shen], [Z].

Remark 8.10. Examples of and [ShY] show that if i) and iii) of Theorem 8.8 are retained but (e.g. if $n = 7$) ii) is weakened to $\mathcal{D}(r,p) = O(r^{1/2})$ then the conclusion fails; see also [AnKLe].

Remark 8.11. For further results related to those of this section; see Shen.

Appendix. Nonnegative Ricci curvature outside a compact set.

Proof of Proposition 8.4: The triangle inequality implies that for $0 < t_1 < t$ the excess function, $\widetilde{E}$, associated to $\gamma(t_1)$, $\gamma(t)$ satisfies

$$E \le \widetilde{E}$$

(where E is the excess function associated to $\gamma(0)$, $\gamma(t)$). Thus it suffices to show that if $\overline{x,p}$ is sufficiently large and t_1, t are suitably chosen, then $\widetilde{E}$ can be made arbitrarily small.

Clearly, we need only ensure that in the present more general situation the bounds on $\Delta G, \Delta\widetilde{E} \mid B_{h(x)}(x)$ are just as in the case in which $\mathrm{Ric}_{M^n} \ge 0$ on all of M^n. This is clear for ΔG, since $B_{h(x)}(x) \cap B_\lambda(p) = \emptyset$, provided $\overline{x,p}$ is sufficiently large.

As for the function $\widetilde{E}$, it clearly suffices to know that a minimal geodesic, σ, from $\gamma(t_1)$ or $\gamma(t)$ to $z \in B_{h(x)}(x)$ does not intersect $B_\lambda(p)$. Consider, for definiteness, the point $\gamma(t_1)$ and suppose $\sigma \cap B_\lambda(p) \ne \emptyset$. Then by the triangle inequality,

$$(*) \qquad \overline{\gamma(t_1), z} > (t_1 - \lambda) + (\overline{x,p} - h(x) - \lambda) \ ,$$

On the other hand, if $\underline{t}$ is a point on γ closest to x, clearly

$$\overline{\gamma(\underline{t}), z} \le 2h(x) \ .$$

Also

$$\underline{t} - t_1 \le \overline{x,p} + h(x) - t_1 \ .$$

Combining these gives

$$(**) \qquad \overline{\gamma(t_1), z} < \overline{x,p} + 3h(x) - t_1 \ .$$

From (*), (**), we get

$$t_1 < 2h(x) + \lambda \ ,$$

and, we can take $t_1 = 2h(x) + \lambda$.

The argument for $\gamma(t)$ is similar to the one just given.

References

[A] U. Abresch: Lower curvature bounds, Toponogov's theorem, and bounded topology I, II, Ann. scient. Ex. Norm. Sup. 18 (1985) 651-670 (4e série), and preprint MPI/SFB 84/41 & 45.

[AG1] U. Abresch and D. Gromoll, "On complete manifolds with nonnegative Ricci curvature, J.A.M.S. (to appear).

[An] M. Anderson, Short geodesics and gravitational instantons, J. Diff. Geom. V. 31 (1990) 265-275.

[AnKle] M. Anderson, P. Kronheimer and C. LeBrun, Complete Ricci-flat Kahler manifolds of inifnite topological type, Commun. Math.Phys. 125 (1989) 637-642.

[Be] M. Berger, Les varietés $\frac{1}{4}$-pincées, Ann. Scuola Norm. Pisa (111) 153 (1960) 161-170.

[BT] R. Bott and L. Tu, Differential forms in algebraic topology, Graduate Texts in Math., Springer-Verlag (1986).

[Ca] E. Calabi, An extension of E. Hopf's maximum principle with an application to Riemannian geometry, Duke Math. J. 25 (1958) 45–56.

[C1] J. Cheeger, Comparison and finiteness theorems in Riemannian geometry, Ph.D. Thesis, Princeton Univ., 1967.

[C2] ———-, The relation between the Laplacian and the diameter for manifolds of nonnegative curvature, Arch. der Math. 19 (1968) 558-560.

[C3] ———-, Finiteness theorems for Riemannian manifolds, Amer. J. Math. 92 (1970) 61-74.

[CE] J. Cheeger and D. G. Ebin, Comparison Theorems in Riemannian Geometry, North-Holland Mathematical Library 9 (1975).

[CGl1] J. Cheeger and D. Gromoll, The splitting theorem for manifolds of nonnegative Ricci curvature, J. Diff. Geo. 6 (1971) 119-128.

[CGl2] ———-, On the structure of complete manifolds of nonnegative curvature, Ann. of Math. 96 (1972) 413-443.

[EH] J. Eschenburg and E. Heintze, An elementary proof of the Cheeger-Gromoll splitting theorem, Ann. Glob. Anal. & Geom. 2 (1984) 141-151.

[GrP] K. Grove and P. Petersen, Bounding homotopy types by geometry, Ann. of Math. 128 (1988) 195-206.

[GrPW] K. Grove, P. Petersen, J. Y. Wu, Controlled topology in geometry, Bull. AMS 20 (2) (1989) 181-

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

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

[GLP] M. Gromov, J. Lafontaine, and P. Pansu, Structure Métrique pour les Variétiés Riemanniennes, Cedic/Fernand Nathan (1981).

[GreWu] R. Greene and H. Wu, Lipschitz convergence of Riemannian manifolds, Pacific J. Math. 131 (1988) 119-141.

[Liu] Z. Liu, Ball covering on manifolds with nonnegative Ricci curvature near infinity, Proc. A.M.S. (to appear).

[M] J. Milnor, Morse Theory, Annals of Math. Studies, Princeton Univ. Press (1963).

[Me] W. Meyer, Toponogov's Theorem and Applications, Lecture Notes, Trieste,1989.

[Pe1] S. Peters, Cheeger's finiteness theorem for diffeomorphism classes of Riemannian manifolds, J. Reine Angew. Math. 394 (1984) 77-82.

[Pe2] ———, Convergence of Riemannian manifolds, Comp. Math. 62 (1987) 3-16.

[ShY] J. P. Sha and D. G. Yang, Examples of manifolds of positive Ricci curvature, J. Diff. Geo. 29 (1) (1989) 95-104.

[Shen] Z. Shen, Finite topological type and vanishing theorems for Riemannian manifolds, Ph.D. Thesis, SUNY, Stony Brook, 1990.

[We] A. Weinstein, On the homotopy type of positively pinched manifolds, Archiv. der Math. 18 (1967) 523-524.

[Z] S. Zhu, Bounding topology by Ricci curvature in dimension three, Ph.D. Thesis, SUNY, Stony Brook, 1990.