CRITICAL POINT THEORY FOR DISTANCE FUNCTIONS

KARSTEN GROVE

INTRODUCTION

One of the fundamental themes in riemannian geometry is to relate properties of a riemannian manifold as a *metric space* to differential *topological* properties of it as a smooth *manifold.*

It is well known that the topology of a smooth manifold is intimately related to the smooth functions it supports via Morse theory (cf. [47]). Applying this fact to functions associated with the manifold as a metric space, such as distance functions, would seem to provide a natural bridge between geometry and topology. The only problem with such a program is that distance functions generally are non smooth, and much less Morse functions. Nonetheless, there is a notion of *regular/critical points* for distance functions, which is *equivalent* to the usual notion for smooth functions (cf. 1.1). The importance of this idea, which was conceived in [37], lies primarily in the observation that some basic principles for smooth functions remain valid for distance functions. In particular, the level set of a regular value is a (topological) submanifold, and the region between two regular levels is (topologically) a product if it contains no critical points (cf. 1.7, 1.8 and 1.14). Moreover, we show that with these techniques a complete Lusternik-Schnirelman theory for distance functions is available (cf. 1.16 - 1.20).

Although the theory has applications in curvature free settings, (cf. e.g. [14], [15], [16], [17], [18], [28], [29], [39], [41], [42] and [71]), it becomes particularly powerful when used in conjunction with Toponogov's comparison theorem, i.e. when a lower curvature bound is present. However, in contrast to earlier work in global

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

riemannian geometry (cf. [20] and [9]), which was based on getting (good) a priori estimates for the injectivity radius, an upper bound for the curvature is irrelevant for this theory. Rather than trying to cover a majority of applications, we will focus on four simple principles, the *convexity* –, *regularity* –, *criticality* –, and *shrinking principle* – (cf. 1.3, 2.5, 2.8 and 2.10). The utility of these principles is then illustrated in different types of problems, namely *recognition* –, *structure* –, and *finiteness problems* (cf. sections 3, 4 and 4).

There are two natural ways in which to extend this theory. One of them is to consider more than one distance function and develop a similar "calculus", in particular an implicit function theorem. Although this has not yet been done systematically, applications of this sort of idea may be found in e.g. [49], [66], and [67]. The other extension is to distance functions on Aleksandrov spaces, i.e. inner metric spaces with a lower curvature bound in the sense of distance comparison (cf. [56]). This extension is actually related to the first. In fact, an "inverse function theorem" for n distance functions on an n-dimensional Aleksandrov space, together with an inverse induction argument form the basis for getting a suitable theory for one function in this setting (cf. forthcoming work of Perelman announced in [6]). – For lack of space and focus, we confine our discussion here to the simplest case of one distance function on a complete riemannian manifold. Strong applications are to be expected, however, when these ideas are used in conjunction with ideas of Gromov-Hausdorff convergence (cf. [30], [33], [34] and [36]), and critical point theory for distance functions on Aleksandrov (limit) spaces.

This paper was written during a visit to Aarhus University. It is a pleasure to thank the department for its hospitality and SNF for its support. Parts of the presentation have been directly influenced by questions and comments from W. Browder , M. Böckstedt, I. Madsen, P. Petersen, J. Tornehare and B. Williams. Also, our discussion of Lusternik-Schnirelman theory owes much to the beautiful expositions given by Palais in [50] and [51]. An apology is due for omitting much of

the beautiful work involving ideas related to those discussed here. In an attempt to make up for this, at least in part, we have included an expanded list of references. At the same time we also refer to the treatments of critical point theory given in the lecture notes by Cheeger [7] and Meyer [46]. – For basic results and tools from riemannian geometry that will be used freely we refer to e.g. [9], [20], [27], [40] or [43].

0. PRELIMINARIES

Throughout M will denote a smooth connected n-dimensional manifold with riemannian metric g. The *distance function*, $\mathrm{dist} : M \times M \to \mathbb{R}$ is defined by

$$\mathrm{dist}(p,q) = \inf \mathrm{length}(c),$$

where the infimum is taken over all piecewise C^1-curves $c : [0,1] \to M$, with $c(0) = p, c(1) = q$ and

$$\mathrm{length}\,(c) = \int_0^1 \sqrt{g(\dot{c}(t), \dot{c}(t))}dt.$$

Ascoli's theorem implies that if (M, dist) is a complete metric space, then any $p, q \in M$ can be joined by a *segment* in M, i.e. a shortest path parametrized proportional to arc length. Curves that are everywhere locally segments are called *geodesics*. Every geodesic is a smooth curve uniquely determined by any one of its velocity vectors. We assume throughout, that (M, dist) is a complete metric space, or equivalently, by the Hopf- Rinow theorem, that every geodesic extends indefinitely in either direction.

For every tangent vector $v \in T_pM \subset TM$ at $p \in M$, let $c_v : \mathbb{R} \to M$ denote the unique geodesic determined by $\dot{c}_v(0) = v$. The *exponential map*, $\exp : TM \to M$ is then defined by

$$\exp\,(v) = c_v(1),$$

and its restriction to T_pM will be denoted by $\exp_p$. A vector $v \in T_pM$ belongs to the *tangent cut locus*, $\mathrm{tancut}(p) \subset T_pM$ if and only if $c_v : [0,1] \to M$ is a segment, but $c_v : [0, 1+\epsilon] \to M$ is not for any $\epsilon > 0$. The *cut locus*, $\mathrm{cut}(p) \subset M$

of p in M is then by definition $\mathrm{cut}(p) = \exp_p(\mathrm{tancut}(p))$. If the *segment domain*, $\mathrm{seg}(p) \subset T_pM$ is the starshaped subset of T_pM bounded by tancut(p), then $\exp_p : \mathrm{seg}(p) \to M$ is surjective and its restriction to $\mathrm{intseg}(p) = \mathrm{seg}(p) - \mathrm{tancut}(p)$ provides a diffeomorphism with $M - \mathrm{cut}(p)$. In particular, the distance function from $p, \mathrm{dist}_p : M \to \mathbb{R}$ is smooth when restricted to $M - \mathrm{cut}(p) \cup p$. Moreover, according to the so-called Gauss lemma, its gradient is the unit radial vector field on this set, i.e. at $q = \exp_p(v) \in M - \mathrm{cut}(p) \cup p$ we have $\mathrm{grad}(\mathrm{dist}_p) = \dot{c}_v(1)/ \parallel \dot{c}_v(1) \parallel \in T_qM$. However, dist_p is clearly not differentiable at points $q \in \mathrm{cut}(p)$ that are joined to p by more than one segment. Nevertheless, it turns out that if these segments do not spread out too much as seen from q, then dist_p behaves as a smooth function near q with $r = \mathrm{dist}(p, q)$ as regular value (cf. section 1).

More generally we need to consider closed subsets $A \subset M$. In this case

$$\mathrm{dist}(A, q) = \min_{p \in A} \mathrm{dist}(p, q)$$

for all $q \in M$. The open and closed r-neighborhoods of A will be denoted by

$$B_A(r) = \{x \in M \mid \mathrm{dist}(A, x) < r\}$$

and

$$D_A(r) = \{x \in M \mid \mathrm{dist}(A, x) \leq r\}$$

respectively. Similarly we set

$$S_A(r) = \{x \in M \mid \mathrm{dist}(A, x) = r\}.$$

As usual, if A is compact, its *diameter* is defined by

$$\mathrm{diam}A = \max_{p \in A} \max_{q \in A} \mathrm{dist}(p, q)$$

whereas its *radius* is

$$\mathrm{rad}A = \min_{p \in A} \max_{q \in A} \mathrm{dist}(p, q).$$

Clearly $\mathrm{rad}A \leq \mathrm{diam}A \leq 2\mathrm{rad}A$ and $A \subset D_p(\mathrm{diam}A)$ for all $p \in A$, whereas $A \subset D_p(radA)$ for some $p \in A$.

1. LUSTERNIK-SCHNIRELMAN THEORY

In this section we will discuss the notion of critical points for distance functions, and show how all of the classical Lusternik-Schnirelman theory for smooth functions on a manifold carries over to this case.

Fix a closed subset A of a complete connected riemannian manifold M. The distance function from A will be denoted by dist_A, i.e.

$$\mathrm{dist}_A : M \to \mathbb{R}, p \to \mathrm{dist}(p, A)$$

for all $p \in M$. As for smooth functions, where the Taylor expansion can be applied, we say that $p \in M$ is a *regular point* for dist_A , or simply for A, if and only if there is a unit vector $v \in S_p \subset T_pM$ and a $c > 0$ such that

$$\mathrm{dist}_A(c_v(t)) \geq \mathrm{dist}_A(c_v(0)) + c \cdot t \tag{1.1}$$

for all sufficiently small $t > 0$. If $p \in M$ is not a regular point for dist_A it is called *critical.*

A simple argument based on standard local distance comparison, shows that definition 1.1 above is equivalent to the following more commonly used characterization. A point $p \in M$ is *regular* for A if and only if there is a $v \in S_p$ such that

$$\angle(v, \dot{c}(0)) > \pi/2 \tag{1.1'}$$

for any segment c from p to A. Similarly, $p \in M$ is a *critical point* for dist_A if and only if every vector $u \in S_p$ makes an angle $\leq \pi/2$ to some segment from p to A.

Thus, whether a point p is regular or critical for A depends entirely on the geometry of segments from p to A, seen from p. More precisely, if $S_{pA} \subset S_p$ denotes the *set of directions* for segments from p to A, then p is *regular* if and only if S_{pA} is contained in an open hemisphere of S_p, and p is *critical* if and only if S_{pA} is a *weak* $\pi/2$-*net* in S_p, i.e. any $u \in S_p$ has distance at most $\pi/2$ to some point of S_{pA}.

Remark 1.2. In a riemannian manifold it is of course equivalent to say that there is a $v \in S_p$ so S_{pA} is contained in the complement of the closed hemisphere centered at v, and to say that there is a $u \in S_p$ such that S_{pA} is contained in the open hemisphere centered at u. This is not the case, however, in singular spaces like orbit - and limit spaces, where the theory works as well with the definition (1.1′) above.

Example 1.3 (Convexity principle) From the work of Cheeger and Gromoll in [10], it is well known that any convex set $C \subset M$ has the structure of a topological manifold with (possibly empty) boundary, ∂C. Moreover, the *interior* int$C$ = $C - \partial C$ is smooth and totally geodesic. If $\partial C \neq \emptyset$, every $p \in \partial C$ has a *supporting halfspace*, i.e. there is a $v \in S_p$ so that any segment from p to an interior point of C makes an obtuse angle with v. In particular, if $A \subset \text{int}C$, then any $p \in \partial C$ is a regular point for A.

Example 1.4 (Cut locus). A point $p \in M - A$ is called a *cut point* for A if and only if no segment from p to A can be extended as a segment to A beyond p. The set of all cut points for A, cut(A) is called the cut locus of A. It is clear from 1.1 that any critical point for A is in cut(A). The converse, however, is usually not true. This in fact, is one of the main reasons for the introduction and utility of the concept.

A simple first variation argument yields the following basic

Lemma 1.5. *Let $\alpha : [0, a] \to M$ be a differentiable curve parametrized by arc length. Suppose there is a $\theta \in [0, \pi/2)$ such that for any $t \in [0, a]$*

$$\angle(\dot{\alpha}(t), v) \geq \pi - \theta$$

for some $v \in S_{\alpha(t)A}$. Then

$$dist_A(\alpha(t)) \geq dist_A(\alpha(0)) + t \cdot \cos\theta,$$

and $\alpha(t)$ is a regular point for each $t \in [0, a]$.

This lemma also shows that if α is a differentiable curve parametrized by arc length, then $S_{\alpha(t)A}$ is completely contained in one of the closed half spaces in $T_{\alpha(t)}M$ determined by $\dot{\alpha}(t)^{\perp}$, unless $\alpha(t)$ is a ***local maximum*** point for $\mathrm{dist}_A \circ \alpha$. So far, however, a reasonable notion of *index* of a critical point is known only in special cases.

In the most important applications of 1.5, α is a curve where $\angle(\dot{\alpha}(t), v) \geq \pi - \theta$ for all $v \in S_{\alpha(t)A}$. Indeed, suppose X is a unit vector field on some open set $U \subset M$ and $\angle(X_p, S_{pA}) \geq \pi - \theta, 0 \leq \theta < \pi/2$, for all $p \in U$. Then any integral curve α of X will satisfy the conditions in 1.5. When allowing θ to depend on $p \in U$, such a *gradient like vector field* exists on all of $U = M - \mathrm{crit}(A)$, where $\mathrm{crit}(A)$ is the closed subset of M consisting of all the critical points for dist_A. This is a direct consequence of 1.1′ and a partition of unity argument. The existence of a fixed θ is of course guaranteed on any compact subset of U.

The first immediate consequence of this discussion is the following important observation, which predates the general idea of critical points as presented here

Proposition 1.6 (Berger Lemma). *Any local maximum point for dist_A is critical.*

In analogy to the case of smooth functions, regular level sets of dist_A have the following structure.

Proposition 1.7 (Implicit Function Theorem). *Let $r > 0$ be a regular value for dist_A, i.e. $\mathrm{dist}_A^{-1}(r) \cap \mathrm{crit}(A) = \emptyset$. Then $\mathrm{dist}_A^{-1}(r) = S_A(r)$ is a topological $n-1$ dimensional manifold, locally flatly embedded in M (cf. also 1.14).*

Proof. Fix a point $p \in \mathrm{dist}_A^{-1}(r)$. Let X be a unit vector field defined in an open neighborhood U of p and satisfying $\angle(X, S_{qA}) \geq \pi - \theta$ for all $q \in U$ and some fixed positive $\theta < \pi/2$. Choose a local hypersurface H through p and transversal to X. For $\epsilon > 0$ sufficiently small, the flow of X defines a diffeomorphism Φ :

$H \times (-\epsilon, \epsilon) \to W \subset M$, where W is an open neighborhood of $p \in M$. By 1.5

$$\mathrm{dist}_A(\Phi(p,\epsilon)) \geq \mathrm{dist}_A(\Phi(p,0)) + \epsilon\cos\theta = r + \epsilon\cos\theta$$

and similarly $\mathrm{dist}_A(\Phi(p,-\epsilon)) \leq r - \epsilon\cos\theta$. Therefore, for H small enough each integral curve $\Phi(q,t), q \in H, t \in (-\epsilon,\epsilon)$ intersects $\mathrm{dist}_A^{-1}(r)$ in exactly one point, $\Phi(q, f(q))$, and $f : H \to (-\epsilon, \epsilon)$ defined this way is obviously continuous. For $\epsilon > 0$ sufficiently small, the map $H \times (-\epsilon,\epsilon) \to M, (q,t) \to \Phi(q, f(q)+t)$ defines the desired submanifold chart for dist_A^{-1} near p. □

As in Morse and Lusternik-Schnirelman theory one has the following key result.

Proposition 1.8 (Isotopy Lemma). *Let $A \subset M$ be a compact subset of M, and suppose $[r_1, r_2] \subset R_+$ contains only regular values for $dist_A$. Then all the levels $dist_A^{-1}(r), r \in [r_1, r_2]$ are homeomorphic, and the annulas*

$$R(r_1,r_2) = D_A(r_2) - B_A(r_1) = \{q \in M | r_1 \leq dist_A(q) \leq r_2\}$$

is homeomorphic to $dist_A^{-1}(r_1) \times [r_1, r_2]$ (cf. also 1.14).

Proof. Compactness of A implies that the sets $D_A(r) = \{q \in M \mid \mathrm{dist}_A(q) \leq r\}, r \geq 0$ are compact. In particular, if X is a gradient like vector field on $M - \mathrm{crit}(A)$, then $\angle(X_p, S_{pA}) \geq \pi - \theta$, for a fixed $0 \leq \theta < \pi/2$ and all p in $R(r_1, r_2)$. It now suffices to invoke 1.5 to complete the proof. □

The proof of 1.8 also applies to the case $r_2 = \infty$, i.e.

Corollary 1.9 (Finite Type Lemma). *Let M be a complete non compact riemannian manifold and $A \subset M$ a compact subset. If $dist_A$ has no critical points in $M - B_A(r)$, then $D_A(r)$ is a compact manifold with boundary $S_A(r)$ and M is diffeomorphic to the interior, $B_A(r)$ of $D_A(r)$.*

In the other extreme case where $r_1 = 0$, we have

Corollary 1.10 (Soul Lemma). *Let $A \subset M$ be a compact submanifold in M without boundary. If there are no critical points in $D_A(r) - A$, then $B_A(r)$ is diffeomorphic to the normal bundle of A in M.*

Proof. First pick $\epsilon > 0$ so that $D_A(\epsilon) \cap \operatorname{cut}(A) = \emptyset$. In particular, there is a unique segment from each $p \in \partial D_A(\epsilon) = \operatorname{dist}_A^{-1}(\epsilon)$ to A. Using this it is easy to construct a gradient like vector field X on $D_A(r) - A$ which is radial near A. The desired diffeomorphism takes each normal segment emanating from A to the corresponding integral curve of X. □

As a trivial combination of 1.9 and 1.10 we get

Theorem 1.11 (Disc Theorem). *Let M be a complete noncompact riemannian n-manifold. If there is a $p \in M$ so that dist_p has no critical points (other than p), then M is diffeomorphic to $\mathbb{R}^n$.*

The compact version of this is contained in

Theorem 1.12 (Sphere Theorem). *Let M be a closed riemannian n-manifold. If there is a $p \in M$ so that dist_p has only one critical point (other than p), then M is homeomorphic to S^n.*

Proof. By assumption, there is only one point q at maximal distance from p, say $d(p, q) = r_0$. Pick $\epsilon > 0$ smaller than the injectivity radii at p and at q. From the isotopy lemma it then follows that $D_p(r)$ is homeomorphic (in fact diffeomorphic, cf. 1.14), to D^n for any $0 < r < r_0$. If therefor $M - B_p(r) \subset B_q(\epsilon)$ for some r, our claim is a consequence of 1.7 and the Generalized Schoenflies theorem (cf. e.g. [57]). Now assume on the contrary, that $M - B_p(r) \not\subset B_q(\epsilon)$ for any $r < r_0$ i.e., there is a sequence of points $x_n \in M$ with $\operatorname{dist}(p, x_n) \to r_0$ and $\operatorname{dist}(q, x_n) \geq \epsilon$ for all n. Since M is compact we find an accumulation point $x \in M$ with $\operatorname{dist}(p, x) = r_0$ and $\operatorname{dist}(x, q) \geq \epsilon$. This contradicts the assumption that only q was critical for dist_p. □

It should be mentioned that any twisted (exotic) sphere has a riemannian metric which satisfies the hypotheses in 1.12. This follows from a general construction due to Weinstein (cf. [4, p. 231]).

In the absence of a good notion of index for critical points, there is nothing to predict the change in topology when crossing a critical level. Rather than pursuing Morse Theory any further, we proceed to show that Lusternik-Schnirelman Theory is valid for distance functions. The key to this is the following

Lemma 1.13 (Deformation Lemma). *Let M be a complete riemannian manifold and $A \subset M$ a compact subset. Suppose $r > 0$ is an isolated critical value of $dist_A$. For every open neighborhood U of $crit(A) \cap dist_A^{-1}(r)$ there is an $\epsilon > 0$, such that $D_A(r+\epsilon) - U$ can be isotoped into $D_A(r-\epsilon)$.*

Proof. As in the proof of 1.7 we see that for each $p \in \mathrm{dist}_A^{-1}(r) - U$, there is an $\epsilon_p > 0$ and a neighborhood U_p of p in M, such that $U_p \cap R(r-\epsilon_p, r+\epsilon_p)$ is homeomorphic to $(U_p \cap \mathrm{dist}_A^{-1}(r)) \times [r-\epsilon_p, r+\epsilon_p]$. By compactness, cover $\mathrm{dist}_A^{-1}(r) - U$ by finitely many sets $U_{p_i} \cap \mathrm{dist}_A^{-1}(r), i = 1, ..., \ell$. With $\epsilon = \min_i \epsilon_{p_i}$, and $W = \bigcup_i U_{p_i}$ clearly $W \cap R(r-\epsilon, r+\epsilon)$ is homeomorphic to $(W \cap \mathrm{dist}_A^{-1}(r)) \times [r-\epsilon, r+\epsilon]$ and $W \cup U \supset \mathrm{dist}_A^{-1}(r)$. By possibly choosing a smaller ϵ we can assume $R(r-\epsilon, r+\epsilon) \subset W \cup U$. In particular $R(r-\epsilon, r+\epsilon) - U \subset W \cap R(r-\epsilon, r+\epsilon)$ and the proof is completed.

□

Remark 1.14. A simple modification of the argument given above shows that the isotopy in 1.13 can be chosen globally on M and fixing everything outside an arbitrarily small neighborhood of $W \cap R(r-\epsilon, r+\epsilon)$. Note also, that if r in 1.13 is a regular value, then U can be chosen empty. By 1.8 and smoothing theory (cf. [44]) it then follows that $\mathrm{dist}_A^{-1}(r)$ has a smooth structure. Moreover, if in 1.8 $\mathrm{dist}_A^{-1}(r)$ is a smooth submanifold for some $r \in [r_1, r_2]$, then all the other levels are smooth as well, and $R(r_1, r_2)$ is diffeomorphic to $\mathrm{dist}_A^{-1}(r) \times [r_1, r_2]$.

The deformation lemma helps in locating critical points other than the obvious minimum and maximum points. The method for this is referred to as the *minimax* *principle.* Let $\mathcal{F}$ be a family of subsets of M. Define the minimax of dist_A over $\mathcal{F}$ by

$$\text{(1.15)} \qquad \mathrm{Minmax}(\mathrm{dist}_A, \mathcal{F}) = \inf_{F \in \mathcal{F}} \sup \{\mathrm{dist}_A(p) | p \in F\}$$

or equivalently

$$\text{(1.15}'\text{)} \qquad \mathrm{Minmax}(\mathrm{dist}_A, \mathcal{F}) = \inf\{r \in \mathbb{R}_+ | \exists F \in \mathcal{F} \,\text{with}\, F \subset D_A(r)\}.$$

A family $\mathcal{F}$ is called *isotopy invariant* if every isotopy of M takes any subset of M from $\mathcal{F}$ to a subset of M from $\mathcal{F}$. From 1.13 (and 1.14) we now get immediately:

Theorem 1.16 (Minimax Principle). *Suppose $A \subset M$ is compact and that $\mathcal{F}$ is an isotopy invariant family of subsets in M. Then Minmax($dist_A, \mathcal{F}$) is a critical value of $dist_A$.*

There are many interesting examples of isotopy invariant families. We mention here only a few of the most important ones.

Examples 1.17 (Isotopy Invariant Families)

(i) Let S be any topological space and $[S, M]$ the set of homotopy classes of maps from S to M. For fixed $f : S \to M$, the family $\mathcal{F}_{[f]} = \{g(S) \subset M | g \in [f]\}$ is clearly an isotopy invariant family.

 (a) If $S = M$ is closed we have $\mathrm{Minmax}(\mathrm{dist}_A, \mathcal{F}_{[id]}) = \mathrm{Max}\ \ \mathrm{dist}_A$.

 (b) If $S = \{\text{point}\}$ then $\mathrm{Minmax}(\mathrm{dist}_A, \mathcal{F}_{[\cdot]}) = \mathrm{Min}\ \ \mathrm{dist}_A = 0$.

 (c) If $S = S^k$, the minimax principle associates to each element $[f] \in \pi_k(M)$ of the k'th *homotopy group* of M a critical value, $\mathrm{Minmax}(\mathrm{dist}_A, \mathcal{F}_{[f]})$ of dist_A.

(ii) Let $H_k(M, R)$ be the k'th singular *homology module* of M with coefficients in a ring R. For each k-cycle z let $\mathcal{F}_{[z]} = \{\text{carrier of } w | w \text{ a } k\text{-cycle with } [w] = [z] \in H_k(M)\}$. Here the carrier of a singular k-chain, c $= \sum n_\alpha \sigma_\alpha, \sigma_\alpha : \Delta_k \to M$ is simply $\bigcup_\alpha \sigma_\alpha(\Delta_k) \subset M$.

(a) If $[M] \in H_n(M)$ is a fundamental class, clearly

Minmax $(\mathrm{dist}_A \mathcal{F}_{[M]}) = \mathrm{Max\ dist}_A$.

(b) If $[\cdot] \in H_0(M)$, then Minmax $(\mathrm{dist}_A,\ \mathcal{F}_{[\cdot]}) = \mathrm{Min\ \ dist}_A = 0$.

(iii) Recall that a subset $X \subset M$ has *Lusternik-Schnirelman category*, $\mathrm{cat}(X; M) = m$ if it can be covered by m (but not fewer) closed subsets of M, each of which is contractible to a point inside M. For each $m \leq \mathrm{cat}(M) := \mathrm{cat}(M; M)$ let $\mathcal{F}_m$ be the family of subsets $X \subset M$ with $\mathrm{cat}(X; M) \geq m$. By the minimax principle

$$c_A(m) = \mathrm{Minmax}(\mathrm{dist}_A, \mathcal{F}_m)$$

is a critical value for each $m \leq \mathrm{cat}(M)$.

Since obviously

$$\mathrm{cat}(X; M) \leq \mathrm{cat}(Y; M) \text{ if } X \subset Y$$

we can also write

$$c_A(m) = \inf\{r \in \mathbb{R}_+ | \mathrm{cat}(D_A(r); M) \geq m\}.$$

From this or $\mathcal{F}_{m+1} \subset \mathcal{F}_m$ we get

$0 = c_A(1) \leq \cdots \leq c_A(m) \leq c_A(m+1) \leq \cdots \leq c_A(\mathrm{cat}(M)) \leq \mathrm{Max\ dist}_A$.

In the last example it can of course happen that equality occurs. For example if $M = \mathrm{RP}^2$ with constant curvature 1 and $A = \{p\}$ then $0 = c_{\{p\}}(1) < c_{\{p\}}(2) = c_{\{p\}}(3) = \pi/2$ (if instead $A = \mathrm{RP}^1 \subset \mathrm{RP}^2$ then $0 = c_{\mathrm{RP}^1}(1) = c_{\mathrm{RP}^1}(2) < c_{\mathrm{RP}^1}(3) = \pi/2$). However, if equality does occur, one gets the following remarkable compensation.

Theorem 1.18 (Main Theorem of L. - S. Theory). *Let M be a complete riemannian manifold and $A \subset M$ a compact subset. For each $m \leq \mathrm{cat}(M)$,*

$$c_A(m) = \mathit{inf}\{r \in \mathbb{R}_+ | \mathrm{cat}(D_A(r); M) \geq m\}$$

is a critical value of $dist_A$ and

$$0 = c_A(1) \le \cdots \le c_A(m) \le \cdots \le c_A(m+k) \le \cdots \le c_A(catM).$$

If moreover, $c_A(m) = c_A(m+k) = c$, then $cat(crit_A(c); M) \ge k+1$ and in particular $\dim crit_A(c) \ge k$.

Proof. It remains to consider the case $c = c_A(m) = c_A(m+k)$. From this and the definition of the c_A's we have

$$\mathrm{cat}(D_A(c-\epsilon); M) \le m-1 \quad \text{and} \quad \mathrm{cat}(D_A(c+\epsilon); M) \ge m+k$$

for any $\epsilon > 0$. Now let U be a neighborhood of $\mathrm{crit}_A(c)$ and choose $\epsilon > 0$ as in 1.13. From trivial properties of $\mathrm{cat}(\cdot, M)$ we then get

$$\begin{aligned}\mathrm{cat}(U; M) &\ge \mathrm{cat}(D_A(c+\epsilon) \cup U; M) - \mathrm{cat}(D_A(c+\epsilon) - U; M)\\ &\ge \mathrm{cat}(D_A(c+\epsilon); M) - \mathrm{cat}(D_A(c-\epsilon); M)\\ &\ge k+1\end{aligned}$$

for every $U \supset \mathrm{crit}_A(c)$. The desired inequality then follows once we have seen that there is a U with $\mathrm{cat}(U; M) = \mathrm{cat}(\mathrm{crit}_A(c); M) = \ell$. For this let $\mathrm{crit}_A(c) \subset F_1 \cup \cdots \cup F_\ell$, where each $F_i, i = 1, \cdots, \ell$ is closed and there are homotopies $\varphi_i : F_i \times [0,1] \to M$ with $\varphi_i(p, 0) = p$ and $\varphi_i(p, 1) = p_i \in M$, for all $p \in F_i, i = 1, \cdots, \ell$. By the homotopy extension property of M we may assume that each φ_i is defined on all of $M \times [0,1]$. For each $i = 1, \cdots, \ell$ let $\mathcal{O}_i$ be a neighborhood of p_i with $\bar{\mathcal{O}}_i$ contractible, and U_i an open neighborhood of F_i with $\bar{U}_i \subset \varphi_i(\cdot, 1)^{-1}(\mathcal{O}_i)$. Then $U = U_1 \cup \cdots \cup U_\ell \supset \mathrm{crit}_A(c)$ has category ℓ because $\mathrm{cat}(U_i; M) = \mathrm{cat}(\bar{U}_i; M) \le \mathrm{cat}(\varphi_i(\bar{U}_i; 1); M) \le \mathrm{cat}(\mathcal{O}_i; M) = \mathrm{cat}(\bar{\mathcal{O}}_i; M) = 1$.

The claim $\dim \mathrm{crit}_A(c) \ge k$ now follows since $\mathrm{cat}(X; M) \le \dim X + 1$ for any closed subset $X \subset M$ (cf. e.g. [50]). □

For any compact subset $A \subset M$ we see in particular that dist_A has at least $\mathrm{cat}(M)$ critical points.

Example 1.19. Let M be the real projective plane RP^2 with riemannian metric so that $\tilde{M}$ is an ellipsoid in $\mathbb{R}^3$ with three different axes. If $A = \{p\}$ is the point in M corresponding to the pair at maximal distance in $\tilde{M}$, then clearly $\mathrm{dist}_A : M \to \mathbb{R}$ has exactly $3 = \mathrm{cat}(\mathrm{RP}^2)$ critical points (including of course $A = p$ itself).

A lower bound for $\mathrm{cat}(M)$ is provided by the so called *cuplength* of M. Here $\mathrm{cuplong}(M)$ is the largest integer ℓ, such that for some field F there are ℓ cohomology classes $\omega_1, \cdots, \omega_\ell \in H^*(M; F)$ each of positive degree and $\omega_1 \cup \cdots \cup \omega_\ell \neq 0$. The following comparison between $\mathrm{cat}(M)$ and $\mathrm{cuplong}(M)$ is proved for metric and path connected spaces M in [5].

Theorem 1.20. *For any riemannian manifold M, cuplong $(M) + 1 \leq$ cat$(M) \leq$* $\dim M + 1$.

This concludes our general discussion of critical point theory for distance functions. In the next sections we will see how to use this in conjunction with comparison theory.

2. COMPARISON THEORY AND CRITICAL POINTS

The utility of critical point theory, as discussed in Section 1, has been particularly apparent so far, in the presence of a lower (sectional) curvature bound. Before we attempt to isolate a few essential ideas behind this, we recall the basic results from comparison theory that are used.

Following [56] we let S^n_k denote the simply connected n-dimensional *space form* of constant curvature k. Points in S^n_k will be written as $\bar{p}, \bar{q}$ etc. rather than p, q etc., which will continue to denote points in a general manifold M.

There are several equivalent formulations of the basic *distance comparison* theorem, usually referred to as Toponogov's triangle comparison theorem. Here are three of them

Theorem 2.1 (Toponogov). *Let M be a complete riemannian manifold with sectional curvature,* $\sec M$ *satisfying* $\sec M \geq k$. *The following equivalent statements* *hold*

(Δ) *For every geodesic triangle* (c_0, c_1, c_2) *in* M *with minimal sides, there is a triangle* $(\bar{c}_0, \bar{c}_1, \bar{c}_2)$ *in* S^2_k *with Length*$(\bar{c}_i) =$ *Length*$(c_i), i = 0, 1, 2$ *and for corresponding angles* $\theta_i \geq \bar{\theta}_i, i = 0, 1, 2$.

(Λ) *Let* $(c_0, c_1; \theta)$ *be any geodesic hinge in* M *with minimal sides, and* $(\bar{c}_0, \bar{c}_i, \theta)$ *the corresponding hinge in* S^2_k. *Then for the hinge endpoints* (p_1, p_0) *and* $(\bar{p}_1, \bar{p}_0)$, *dist*$(p_1, p_2) \leq$ *dist*$(\bar{p}_1, \bar{p}_0)$.

(T) *Consider any pair* $(c_0; p_0)$, *where* c_0 *is a minimal geodesic in* M *and* $p_0 \in M$. *Let* $(\bar{c}_0; \bar{p}_0)$ *be the corresponding pair in* S^2_k, *i.e. the distances from* $\bar{p}_0$ *to the endpoints of* $\bar{c}_0$ *are the same as from* p_0 *to the endpoints of* c_0. *Then dist*$(p_0, q) \geq$ *dist*$(\bar{p}_0, \bar{q})$ *for any* $q \in c_0$ *and corresponding* $\bar{q} \in \bar{c}_0$.

In each of these statements c_0 *does not need to be minimal, only Length*$(c_0) \leq \pi/\sqrt{k}$ *if* $k > 0$. *In this case, however, the angle comparison in* (Δ) *holds only for the angles adjacent to* c_0.

There are important rigidity companions to (T) *and* (Λ) *above in cases of equality:*

[Λ] *Suppose* $0 < \theta < \pi$ *and dist*$(p_1, p_0) =$ *dist*$(\bar{p}_1, \bar{p}_0)$. *Then* (c_0, c_1) *spans a surface in* M *isometric to the unique triangular surface in* S^2_k *spanned by* $(\bar{c}_0, \bar{c}_1)$, *and with totally geodesic interior.*

[T] *Assume* $p_0 \notin c_0$ *and dist*$(p_0, q) =$ *dist*$(\bar{p}_0, \bar{q})$ *for some interior* $q \in c_0$. *Then every minimal geodesic* c_q *from* p_0 *to* q *spans together with* c_0 *a unique surface isometric to the triangular surface in* S^2_k *spanned by* $\bar{c}_0$ *and* $\bar{p}_0$, *and with totally geodesic interior.*

One advantage with the T-version of 2.1 is that it makes sense in more general inner metric spaces where angles are not a priori defined (cf. [56]). If S is such a space with curv$S \geq k$, i.e. local distance comparison ála 2.1 holds in S, then indeed global distance comparison holds as well (see [6]). The corresponding rigidity results have been proved and applied recently in [30].

It is sometimes useful to interpret (Λ) in terms of the exponential maps as in [34]: If $p \in M$ and $\mathrm{rad}(p) = \mathrm{maxdist}_p$ we endow $D_0(\mathrm{rad}(p)) \subset T_pM$ with the constant curvature k metric obtained from the euclidean metric by a radial conformal change (when $k > 0$ and $\mathrm{rad}(p) = \pi/\sqrt{k}$ we interpret $D_0(\mathrm{rad}(p))$ as S^n_k). In this way we view the segment domain $\mathrm{seg}(p)$ as a subset of S^n_k. Clearly 2.1(Λ) is equivalent to

$$(2.1') \qquad \exp_p : \mathrm{seg}(p) \to M \quad \text{is distance nonincreasing.}$$

This is the basis also for volume comparison of various important *metrically defined subsets* of M. We mention only two examples of this, both of which are special cases of one general result from [33] (cf. also [11]).

Example 2.2 (Half spaces). Fix $p \in M$ and a closed subset $Q \subset M$. Let $\bar{p} = \exp_p^{-1}(p)$ and $\bar{Q} = \exp_p^{-1}(Q)$ in $\mathrm{seg}(p) \subset S^n_k$. For the *half-spaces* $H(p,Q) = \{x \in M | \mathrm{dist}(x,p) \le \mathrm{dist}(x,Q)\}$ and $H(\bar{p},\bar{Q})$ in M and S^n_k respectively we then have

$$\mathrm{vol}\, H(p,Q) \le \mathrm{vol}\, H(\bar{p},\bar{Q}).$$

Example 2.3 (Swiss Cheeses). With p and Q as in 2.2 fix $R > 0$ and an arbitrary function $r : Q \to \mathbb{R}_+$. By definition the *swiss cheese* $K = K((Q,r);(p,R))$ in $D_p(R)$ relative to r is the set $K = D_p(R) - \cup_{q \in Q} B_q(r(q))$. Then

$$\mathrm{vol}\, K((Q,r);(p,R)) \le \mathrm{vol}\, K((\bar{Q},\bar{r});(\bar{p},R))$$

where $\bar{r} = r \circ \exp_p : \bar{Q} \to \mathbb{R}_+$.

The volume estimates given in 2.2 and 2.3 <u>do not hold</u> under the weaker curvature assumption Ric $M \ge (n-1)k$. This, however, is sufficient for the following simple extension of the so-called Bishop-Gromov volume comparison theorem (cf. e.g. [7]).

Theorem 2.4 (Relative Volume Comparison). *Let M be a complete riemannian n-manifold with Ric $M \geq (n-1)k$ and suppose $Q \subset M$ is compact. If we set $v_k^n(R) =$ vol $D_{\bar{p}}(R)$ in S_k^n, then*

$$R \to vol\, D_Q(R)/v_k^n(R)$$

is a nonincreasing function.

We are now ready to present some principles frequently used in the detection of critical or regular points.

Throughout we fix a complete riemannian n-manifold M with $\sec M \geq k$.

In the *regularity principle* we consider two points $p, q \in M$ and fix $\bar{p}, \bar{q} \in S_k^2$ with $\operatorname{dist}(p,q) = \operatorname{dist}(\bar{p},\bar{q})$. Except for the single case where $k > 0, d(p,q) = \pi/\sqrt{k}$ and in particular $M \equiv S_k^n$, there is a continuous map $T : M \to S_k^2$ defined by the requirement $\operatorname{dist}(p,x) = \operatorname{dist}(\bar{p}, T(x))$ and $\operatorname{dist}(q,x) = \operatorname{dist}(\bar{q}, T(x))$ for all $x \in M$. T is unique up to reflection in the segment, $\bar{p}\bar{q}$ from $\bar{p}$ to $\bar{q}$ in S_k^2. If

$$\operatorname{reg}(\bar{p},\bar{q}) = \{\bar{x} \in S_k^2 | \angle(\bar{p},\bar{x},\bar{q}) > \pi/2\}$$

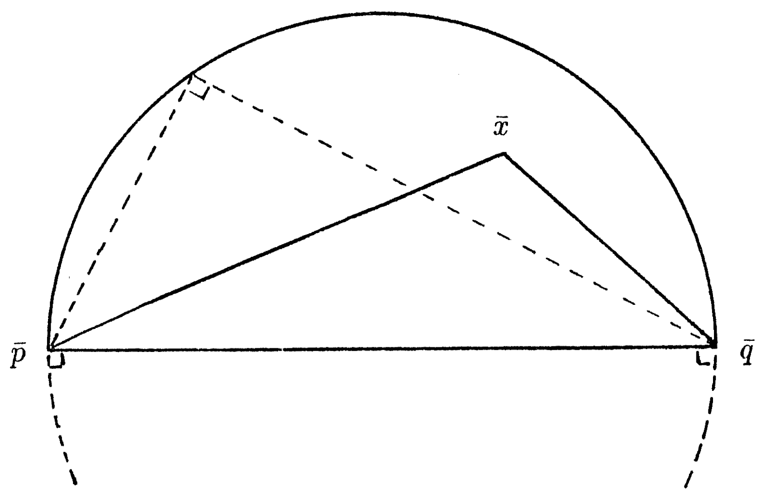
(cf. figure 2.6) we have

Lemma 2.5 (Regularity Principle). *For any $p, q \in M$ the set*

$$\operatorname{reg}(p,q) := T^{-1}(\operatorname{reg}(\bar{p},\bar{q}))$$

consists of regular points for p as well as for q.

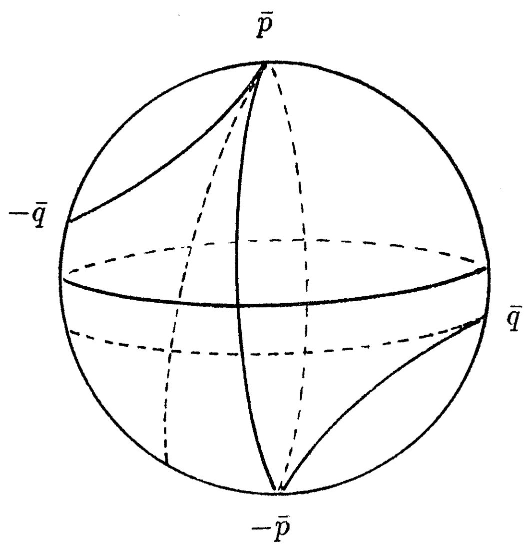
Proof. Let $x \in T^{-1}(\operatorname{reg}(\bar{p},\bar{q}))$. Then by 2.1($\Delta$), the angle at x between any two segments from x to p and from x to q is obtuse. $\square$



reg $(\bar{p}, \bar{q})$ for $k = 0$

FIGURE 2.6

In the special case where $k > 0$ and $\mathrm{dist}(\bar{p}, \bar{q}) = \pi/2\sqrt{k}, \mathrm{reg}(\bar{p}, \bar{q}) \subset S_k^2$ has two connected components, one of which is $B(\bar{p}, \pi/2\sqrt{k}) \cap B(\bar{q}, \pi/2\sqrt{k})$. If $\mathrm{dist}(\bar{p}, \bar{q}) = D > \pi/2\sqrt{k}$ then $S_k^2 - \overline{\mathrm{seg}(\bar{p}, \bar{q})} = \mathrm{seg}(-\bar{p}, \bar{q}) \cup \mathrm{seg}(-\bar{q}, \bar{p})$, and in particular $\mathrm{seg}(\bar{p}, \bar{q}) \supset B_{\bar{p}}(D) \cap B_{\bar{q}}(D)$ (cf. figure 2.7).



reg $(\bar{p}, \bar{q})$ for $k > 0$ and dist $(\bar{p}, \bar{q}) > \pi/2\sqrt{k}$

FIGURE 2.7

In the *criticality principle* we consider points $p, q \in M$ where $q \in \text{crit}(p)$. Fix corresponding points $\bar{p}, \bar{q} \in S_k^2$ and for each $r > \text{dist}(\bar{p}, \bar{q}) = \text{dist}(p, q) = d$ let $\bar{r} \in S_k^2$ be the unique point (up to reflection in segment $\bar{p}\bar{q}$) with $\text{dist}(\bar{p}, \bar{r}) = r$ and $\angle(\bar{p}, \bar{q}, \bar{r}) = \pi/2$ (cf. figure 2.9 below). If $\theta_k(r, d) = \angle(\bar{r}, \bar{p}, \bar{q})$ then $\theta_k(r, d)$ is increasing in r and decreasing in d (if $k > 0$ we assume here $d < r < \pi/2\sqrt{k}$). With this notation:

Lemma 2.8 (Criticality Principle). *Let $p \in M$ and suppose $q \in crit(p)$ with $dist(p, q) = d$. Then for all $x \in M$ with $dist(p, x) = r > d$, the angle between any pair of segments from p to q and from p to x is at least $\theta_k(r, d)$.*

Proof. Let c_0 be a segment from p to q, c_1 a segment from p to x and c_2 a segment from q to x. If $\angle(\dot{c}_0(0), \dot{c}_1(0)) < \theta_k(r, d)$ so is the comparison angle in the triangle $(\bar{c}_0, \bar{c}_1, \bar{c}_2)$ by 2.1(Δ). By definition of $\theta_k(r, d)$, therefore, $\angle(-\dot{\bar{c}}_0(d), \dot{\bar{c}}_2(0)) > \pi/2$. This means, however, that also $\angle(-\dot{c}_0(d), \dot{c}_2(0)) > \pi/2$ by 2.1(Δ). Since c_0 was arbitrary, this contradicts the assumption $q \in \text{crit}(p)$. Thus $\angle(\dot{c}_0(0), \dot{c}_1(0)) \geq \theta_k(r, d)$ as claimed. $\square$

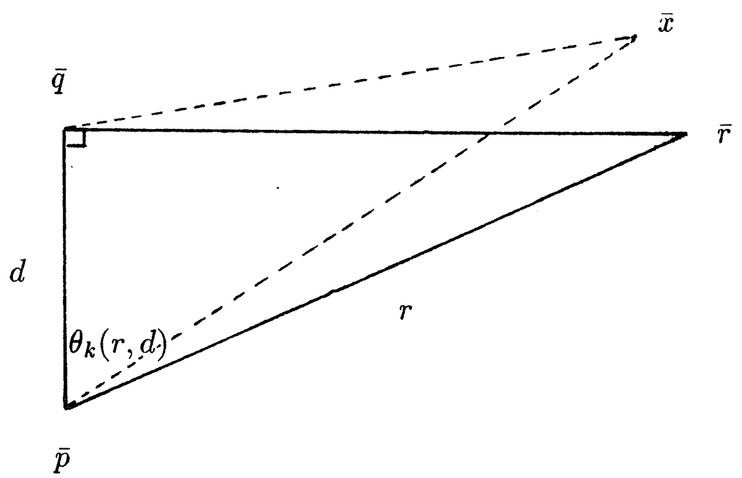


FIGURE 2.9

The *shrinking principle* below is based on 2.1′ as applied in the volume estimates 2.2 and 2.3.

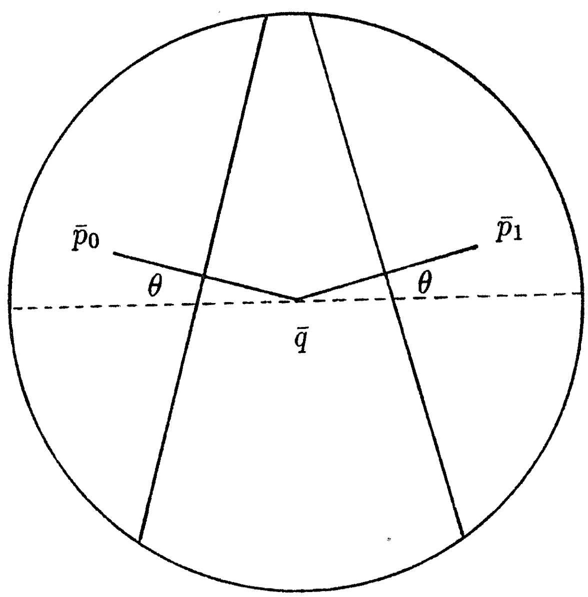
Lemma 2.10 (Shrinking Principle). *Let $p \in M$ and suppose $q \in crit(p)$ with $dist(p,q) = d$. Then*

(i) *$vol\ H(q;p) \leq vol\, H(\bar{q}, \{\bar{p}_0, \bar{p}_1\})$, where $H(\bar{q}, \{\bar{p}_0, \bar{p}_2\}) \subset D_q(\mathrm{rad}(q)) \subset S^n_k$, $\bar{q}$ is the midpoint of segment $\bar{p}_0\bar{p}_1$ and $dist(\bar{q}, \bar{p}_i) = d$.*

(ii) *$vol(M - B_p(r)) \leq vol(D_{\bar{q}}(\mathrm{rad}(q)) - (B_{\bar{p}_0}(r) \cup B_{\bar{p}_1}(r)))$ for any $r > 0$.*

Proof. From 2.2 and 2.3 we have $\mathrm{vol}\, H(q;p) \leq \mathrm{vol}\, H(\bar{q}, \exp_{\bar{q}}^{-1}(p))$ and $\mathrm{vol}(M - B_p(r) \leq \mathrm{vol}\, K((\exp_{\bar{q}}^{-1}(p), r); (\bar{q}, \mathrm{rad}(q)))$. It remains to see that if we replace $\exp_{\bar{q}}^{-1}(p)$ by $\{\bar{p}_0, \bar{p}_1\}$ chosen as in 2.10, then we get subsets in S^n_k with (possibly) even bigger volume. For this observe that $S_{qp} \subset T_qM$ forms a weak $\pi/2$ – net in the unit sphere $S_q \subset T_qM$ because $q \in \mathrm{crit}(p)$. From 2.4 we conclude that vol $D_{S_{qp}}(\theta)/v_1^{n-1}(\theta) \geq$ vol $S_{qp}(\pi/2)/v_1^{n-1}(\pi/2) = \mathrm{vol}\, S_1^{n-1}/v_1^{n-1}(\pi/2) = 2$ for all $\theta \leq \pi/2$, where $D_{S_{qp}}(\theta) \subset S_q$ is the θ-neighborhood of S_{qp} in $S_q \equiv S_1^{n-1}$. Thus $\mathrm{vol}\, D_{S_{qp}}(\theta) \geq \mathrm{vol}\, D_{\{v_0, v_1\}}(\theta)$, where $v_0, v_1 \in S_q$ is any antipodal pair. By integrating the reverse inequality for the corresponding complements we derive our claim.

□



$H(\bar{q}; \{\bar{p}_0, \bar{p}_1\}) \subset D_{\bar{q}}(R)$

FIGURE 2.11

Remark 2.12. The shrinking lemma has a natural generalization, where instead of assuming q to be a critical point for p, one assumes that S_{qp} is a weak $\frac{1}{2}\pi + \theta$ net in S_q for some $0 \leq \theta < \pi/2$. In this case the points $\bar{p}_0, \bar{p}_1$ must be chosen so that $\angle(\bar{p}_0\bar{q}\bar{p}_1) = \pi - 2\theta$, where the corresponding $\{v_0, v_1\}$ form a weak $\frac{1}{2}\pi + \theta$ net in S_q.

In the subsequent sections we will apply these principles to different types of problems in riemannian geometry.

3. RECOGNITION THEOREMS

In this section we give examples of situations where critical point theory is used to determine the type of the manifold. The simplest and first such application was given in [36]:

Theorem 3.1 (Diameter Sphere Theorem). *Any complete riemannian manifold M with $secM \geq 1$ and $diam\, M > \pi/2$ is a twisted sphere.*

Proof. Choose $p, q \in M$ with $\text{dist}(p, q) = \text{diam}\, M$. By the regularity principle 2.5 all points in $M - \{p, q\}$ are regular for p as well as for q (cf. also fig. 2.7). The conclusion then follows from 1.12. $\square$

We point out that in the special case of 3.1, one does not have to appeal to the generalized Schoenflies Theorem as we did in 1.12. It suffices to observe that a gradient like vector field X can be constructed on $M - \{p, q\}$ which is radial near p and q. We also note that the following question remains open.

Problem 3.2. *Is there a $d = d(n) < \pi$ so that any complete riemannian n-manifold M with $\sec M \geq 1$ and $diam\, M \geq d$ is diffeomorphic with S^n?*

If in this problem we replace the diameter by the radius, an affirmative answer has been given in [65]. The idea here is to show that when $\sec M \geq 1$ and $\text{rad}M \sim \pi$, then M is metrically close to S_1^n in the Gromov-Hausdorff, and hence Lipschitz sense (cf. [70]).

Except for the question about existence of different differentiable structures on manifolds M as in 3.1, it is well known that the diameter sphere theorem is optimal. However, when $\operatorname{diam} M > \pi/2$ is replaced by $\operatorname{diam} M \geq \pi/2$ one has the following essentially complete metric classification proved in [21] and [22].

Theorem 3.3 (Diameter Rigidity). *Let M be a complete riemannian manifold with $secM \geq 1$ and diam $M = \pi/2$. Then, either M is a twisted sphere or else*

(i) *If $\pi_1(M) = \Gamma \neq \{1\}$, M is isometric to*

  (a) *The unique $\mathbb{Z}_2$-quotient of a complex odd dimensional projective space, or*

  (b) *S_1^n/Γ, where $\Gamma \to 0(n+1)$ is reducible.*

(ii) *If $\pi_1(M) = \{1\}$, M is isometric to a projective space, except possibly if M has the cohomology ring of the Cayley plane, C_aP^2.*

In this theorem, one would also have rigidity in the exceptional case where $H^*(M) \cong H^*(C_aP^2)$, provided the following holds

Conjecture 3.4. *Any riemannian submersion $S_1^{15} \to M^8$ is congruent to the Hopf map $S_1^{15} \to S_4^8$.*

The proof of 3.3 is rather long and intricate. Here we only give an outline in order to show how critical point theory enters:

As in 3.1 we begin by choosing points $p, q \in M$ with $\operatorname{dist}(p,q) = \operatorname{diam} M = \pi/2$. Then

$$A = \{x \in M | \operatorname{dist}_p(x) = \pi/2\} \text{ and } \bar{A} = \{x \in M | \operatorname{dist}_A(x) = \pi/2\}$$

are both non-empty, and totally π-convex by 2.1 (T), i.e. any geodesic in M of length $< \pi$ and with endpoints in A (resp. $\bar{A}$) is entirely contained in A (resp. $\bar{A}$).

The regularity principle 2.5 shows that all points in $M - (A \cup \bar{A})$ are regular for dist_A as well for $\operatorname{dist}_{\bar{A}}$. In particular, for any $\epsilon > 0$, $M - B_A(\epsilon) \cup B_{\bar{A}}(\epsilon)$ is homeomorphic to $\operatorname{dist}_A^{-1}(\epsilon) \times [0,1]$. Moreover, from the structure of convex sets (cf. 1.3) A(resp.$\bar{A}$) is either a closed smooth totally geodesic submanifold of M or else a compact topological submanifold with $\partial A \neq \emptyset$ and smooth totally geodesic interior. In the first case, $B_A(\epsilon)$ is diffeomorphic to the normal bundle of A in M by 1.10. If on the other hand $\partial A \neq \emptyset$, then $\mathrm{dist}_{\partial A} : A \to \mathbb{R}$ is strictly concave since $\sec M > 0$ (cf. [10]). In particular there is a unique point $s \in \mathrm{int} A$ at maximal distance from ∂A and all points in A are s-regular as explained in 1.3. Thus for $\epsilon > 0$ sufficiently small, all points in $D_A(\epsilon)$ are regular for dist_s and therefor $D_A(\epsilon)$ is diffeomorphic to the $\dim M$ dimensional euclidean disc by 1.10. – All in all we conclude that there are totally geodesic smooth submanifolds of M (possibly points), so that M is the union of their tubular neighborhoods.

Observe in particular, that if $\partial A \neq \emptyset$, and $\partial \bar{A} \neq \emptyset$ then M is a twisted sphere.

Now suppose M is not a sphere. We claim that $\partial A = \partial \bar{A} = \emptyset$ i.e., A and $\bar{A}$ are smooth totally geodesic submanifolds of M (one of them possibly a point). Indeed, if say $\partial \bar{A} \neq \emptyset$ pick $\bar{p} \in \bar{A}$ arbitrarily and let c be a segment from $\bar{p}$ to A. By the rigidity comparison theorem 2.1 [Λ], it follows that any normal vector $u \in T_q A^{\perp}$ which is obtained from $-\dot{c}(\pi/2) \in T_{c(\pi/2)} A^{\perp}$ by parallel translation along a curve in A, defines a segment c_u from $q \in A$ to $\bar{p}$. Moreover, this set $U \subset TA^{\perp}$ of normal vectors to A is a smooth closed submanifold of the unit normal bundle $SA^{\perp}$ to A with fiber $F \subset U$ being the orbit of $-\dot{c}(\pi/2)$ in $S_{c(\pi/2)} A^{\perp}$ under the normal holonomy group. By assumption $SA^{\perp} \simeq \mathrm{dist}_A^{-1}(\epsilon) \simeq \mathrm{dist}_{\bar{A}}^{-1}(\epsilon)$ is diffeomorphic to the $\dim M - 1$ dimensional sphere. If therefore $SA^{\perp} - U \neq \emptyset$, the (normal) projection $U \to A$ is homotopic to a constant map $U \to pt \in A$. Using the homotopy lifting property of $U \to A$ we get a homotopy $h : U \times [0,1] \to U$ where $h_0 = id_U$ and $h_1(U) \subset F$. Since U and $F \subset U$ are closed manifolds this is impossible and hence $U = SA^{\perp}$. This on the other hand implies that every normal vector to A defines a segment from A to $\bar{p} \in \bar{A}$. Since $\bar{p} \in \bar{A}$ was chosen arbitrarily this yields the desired contradiction.

An elaboration of the argument just given shows that $\mathrm{cut}(A) = \bar{A}$ and $\mathrm{cut}(\bar{A}) =$

A. Moreover, for each $\bar{p} \in \bar{A}$ the map

$$S_{\bar{p}}\bar{A}^{\perp} \to A,\ \bar{v} \to \exp_{\bar{p}}(\pi/2 \cdot \bar{v})$$

is a riemannian submersion from the euclidean unit normal sphere, $S_{\bar{p}}\bar{A}^{\perp}$ onto A. An essentially complete metric classification of such fibrations was given in [22]. The only case left is when $S_{\bar{p}}\bar{A}^{\perp} \equiv S_1^{15}$ and A is a simply connected 8-dimensional manifold $\simeq S^8$ (cf. 3.4).

The remaining part of the proof of 3.3 is separated into the cases (i) $\pi_1(M) = \Gamma \neq \{1\}$ and (ii) $\pi_1(M) = \{1\}$. – The topological decomposition of M obtained via critical point theory is used together with Morse theory for geodesics to show that when $\pi_1(M)$ is trivial so are $\pi_1(A)$ and $\pi_1(\bar{A})$. In this case, the classification of fibrations $S_{\bar{p}}\bar{A}^{\perp} \to A$ gives in particular that A and $\bar{A}$ are rank 1 symmetric spaces and then that M itself is such a space. – When $\pi_1(M) \neq \{1\}$, one considers the universal cover $\tilde{M}$ of M. Clearly sec $\tilde{M} \geq 1$ and diam $\tilde{M} \geq \pi/2$. In view of the classification given already for simply connected manifolds the remaining case of interest is when diam $\tilde{M} > \pi/2$. In particular, $\tilde{M}$ is a topological sphere by 3.1, which is decomposed similarly using the lifts $\tilde{A}, \tilde{\bar{A}}$ of $A, \bar{A}$. By a second variation argument of Synge type we have in general dim $A +$dim $\bar{A} \leq$ dim $M - 1$. However, when $\tilde{M}$ is a sphere a simple transversality argument then implies dim $\tilde{A} +$dim $\tilde{\bar{A}} =$ dim $\tilde{M} - 1$. The riemannian fibrations constructed above are then local isometries. It is now fairly easy to show that $\tilde{A}, \tilde{\bar{A}}$ is an orthogonal pair of totally geodesic subspheres in the unit sphere $\tilde{M}$. This concludes the outline of 3.3. $\square$

Before leaving the class of manifolds M with sec $M \geq 1$ and diam $M \geq \pi/2$ we like to point out some interesting volume problems related to critical point theory. The first is a natural analogue of a classical area problem due to A. D. Aleksandrov (cf. [34])

Conjecture 3.5. *Let M be a closed riemannian n-manifold with* sec $M \geq 1$ *and diam $M = d > \pi/2$. Then vol$M < 2v_1^n(d/2)$, and this estimate is optimal. Note* *that* vol$X$ $= 2\ v_1^n(d/2)$ *for the singular spherical space* X *obtained by gluing to copies of* $D_1^n(d/2) \subset S_1^n$ *together along their boundary.*

If in 3.5 we replace the diameter by the radius an optimal estimate has been found in [34]. There it was also proved that if sec $M \geq 1$ and rad $M > \pi/2$, then every dist_p, $p \in M$ has exactly two critical points (including p itself). This then gives a lower bound for the Filling Radius, and hence the volume of M (cf. [24]). The optimal lower bound, however, is not known. The following was proposed in [34].

Conjecture 3.6. *If* M *is a closed riemannian* n*-manifold with* sec $M \geq 1$ *and rad* $M = r > \pi/2$, *then* vol$M \geq$ vol$S^n_{(\pi/r)^2}$, *i.e., the volume of* M *is at least the volume of the constant curvature* n*-sphere with diameter* r.

We close this section with a recognition theorem for exotic spheres. Following the terminology of [3], the *excess function* associated with $p, q \in M$ is given by

$$\mathrm{exc}_{p,q}(x) = \mathrm{dist}(p,x) + \mathrm{dist}(x,q) - \mathrm{dist}(p,q) \tag{3.7}$$

for all $x \in M$. Based on this we define the *excess* of M, exc$M$ as in [34] by

$$\mathrm{exc}M = \min_{p,q} \max_{x} \mathrm{exc}_{p,q}(x). \tag{3.8}$$

Other excess type invariants have been introduced in [48], [59] and [30]. Observe that exc$M = 0$ if and only if there are points $p, q \in M$ such that $\mathrm{cut}(p) = \{q\}$ and $\mathrm{cut}(q) = \{p\}$. In particular M is a twisted sphere. Conversely, any twisted sphere has a riemannian metric with excess $= 0$ (cf. [4,p.231]). However, it is easy to see that small excess has no topological significance in general: Simply take any manifold M and concentrate all of its topology in a tiny metric ball whose complement is the complement of a tiny ball in the unit sphere. In this simple construction there is of course no curvature control. The following problem posed in [35] appears to be significant.

Problem 3.9. *For fixed $k \in \mathbb{R}$ and $D \in \mathbb{R}_+$ describe all closed riemannian n-manifolds with* sec $M \geq k$, *diam* $\leq D$ *and exc$M$ arbitrarily small.*

By excluding the possibility of collapse one has the following answer from [35].

Theorem 3.10 (Exotic Sphere Theorem). *Given an integer $n \geq 2$, a real k and $v, D > 0$. There is an $\epsilon = \epsilon(n, k, D, v)$ such that any closed riemannian n-manifold M with* sec $M \geq k$, *diam$M \leq D$ and vol$M \geq v$ is a homotopy sphere whenever* exc $M \leq \epsilon$.

Proof. Pick $p, q \in M$ with max $\mathrm{exc}_{p,q} = \mathrm{exc}\, M$. A simple application of the regularity principle 2.5 shows that for every $\delta > 0$ there is an $\epsilon = \epsilon(\delta, k, D)$ so that $M - B_p(\delta) \cup B_q(\delta)$ consists of regular points for p as well as for q, whenever $\mathrm{exc}\, M \leq \epsilon$. In view of the isotopy lemma 1.8, therefore, it suffices to show that $B_p(\delta)$ and $B_q(\delta)$ are contractible to points inside M: Indeed, in this case $M = X_1 \cup X_2$, where $X_1 = M - B_q(\delta)$ and $X_2 = M - B_p(\delta)$ are contractible in M. Then $H^*(M, X_i) \to H^*(M)$ is surjective for $i = 1, 2$ and therefor $H^*(M)$ has trivial cup product structure for any coefficient ring. Using $\mathbb{Z}_2$ as coefficient field, it follows from Poincaré duality that, in particular $H_1(M; \mathbb{Z}_2) = 0$. From the Mayer-Victoris sequence

$$0 \to H_0(X_1 \cap X_2; \mathbb{Z}_2) \to H_0(X_1; \mathbb{Z}_2) \oplus H_0(X_2; \mathbb{Z}_2) \to H_0(M; \mathbb{Z}_2) \to 0$$

we conclude that $X_1 \cap X_2$ is connected. This in turn implies that $\pi_1(M) = \{1\}$ by Van Kampen's theorem. Now any simply connected homology sphere is a homotopy sphere by theorems of Hurewicz and Whitehead. The proof of 3.10 is therefor complete once the following basic *local geometric contractibility* result has been established. □

Theorem 3.11 (LGC-Lemma). *Given an integer $n \geq 2$, a real k and $D, v > 0$. There is a $\delta = \delta(n, k, D, v)$ such all points $(p, q) \in M \times M$ with dist$(p, q) \leq \delta$ are regular for the diagonal $\Delta(M) \subset M \times M$. In particular, for each $p \in M, D_p(\delta)$ is contractible in M to a point.*

Proof. Assume $(p,q) \in M \times M$ is Δ-critical and $\mathrm{dist}(p,q) = d$. Then clearly p is q-critical and q is p-critical. By the shrinking principle 2.10, $\mathrm{vol}\, M = \mathrm{vol}\, H(p;q) + \mathrm{vol}\, H(q;p) \leq 2\,\mathrm{vol}\, H(\bar{q};\{\bar{p}_0,\bar{p}_1\})$, where $H(\bar{q};\{\bar{p}_0,\bar{p}_1\}) = \{\bar{x} \in D_{\bar{q}}(D) | \mathrm{dist}(\bar{q},\bar{x}) \leq \mathrm{dist}(\bar{x},\{\bar{p}_0,\bar{p}_1\})\} \subset S^n_k$ and $\bar{q}$ is the midpoint of $\bar{p}_0\bar{p}_1$ and $\mathrm{dist}(\bar{q},\bar{p}_i) = d$ (cf. fig. 2.11 with $\theta = 0$). Since obviously $\mathrm{vol}\, H(\bar{q};\{\bar{p}_0,\bar{p}_1\}) \to 0$ and $d \to 0$ this proves the first claim. – The deformation retraction defined near $\Delta \subset M \times M$ by following the integral curves of a gradient like vector field for dist_Δ, will also provide the desired deformation of $D_p(\delta)$ via the embedding $D_p(\delta) \subset \{p\} \times D_p(\delta) \subset M \times M$. □

Remark 3.12 (Contractibility functions). By appealing to the more general θ-version 2.12 of the shrinking lemma, one gets an important sharpening of the LGC-lemma above: There are $\delta = \delta(n,k,D,v) > 0$ and $\theta = \theta(n,k,D,v) > 0$ such that any point $(p,q) \in M \times M$ with $\mathrm{dist}(p,q) \leq \delta$ is *θ-regular*, i.e. $S_{(p,q)\Delta}$ is contained in a $\frac{1}{2}\pi - \theta$ ball in $S_{(p,q)} M \times M$. From 1.5 we then find an $R = R(n,k,D,v)$ so that any r-ball in M with $r \leq \delta$ is contractible inside the concentric ball of radius $R \cdot r$.

The local geometric contractibility control described above is crucial for example in the derivation of homotopy - and homeomorphism finiteness results (cf. section 5).

Remark 3.13. An application of the relative volume comparison theorem 2.4 shows that any n-manifold M with $\mathrm{Ric}\, M \geq (n-1)$ and $\mathrm{diam}\, M \geq \pi - \epsilon$ has small excess. Theorem 3.10 therefor has a diameter Ricci curvature sphere theorem as corollary (cf. [35]).

4. STRUCTURE THEOREMS

It is too optimistic to expect a recognition type solution to the excess problem mentioned in 3.9. Rather, one would hope to at least be able to find restrictions for the structure of such manifolds. The applications of critical point theory given in this section are to problems of that kind.

Although all manifolds M in this section will be complete and non compact, the idea in the following observation due to Gromov plays a key role in the Betti-number finiteness theorm of section 5.

Theorem 4.1. *Let M be a non compact riemannian manifold with* $\sec M \geq 0$. *Then M is diffeomorphic to the interior of a compact manifold with boundary.*

Proof. We claim that for any $p \in M$, there is an $r > 0$ so that dist_p has no critical points in $M - B_p(r)$. In fact, if this is not the case let $q_i, i = 1, 2, \ldots$ be a sequence of points in $\mathrm{crit}(p)$ with say $\mathrm{dist}(p, q_{n+1}) \geq 2\,\mathrm{dist}(p, q_n)$. By the criticality principle 2.8 the angle between any segment from p to q_i and from p to q_j, is bounded below by $\theta_0(1,2)(= \theta_0(d, 2d)$ for any $d > 0)$ independent of i, j. This is of course impossible since $S_p \subset T_pM$ is compact. The claim now follows from the finite type lemma 1.9. □

The following much stronger structure theorem is due to Cheeger and Gromoll [10]:

Theorem 4.2 (Soul Theorem). *Any complete non compact riemannian manifold M with* $\sec M \geq 0$, *is diffeomorphic to the normal-bundle of some compact totally geodesic submanifold $S \subset M$.*

The dominating feature in the proof of this result is convexity. Here is an outline:

Choose $p \in M$ arbitrarily and let $c : [0, \infty] \to M$ be a ray in M, i.e. $c|[s,t]$ is a segment for any s,t. The existence of such a ray is a simple consequence of the Hopf-Rinow theorem. For each $t \geq 0$, $H_t(c) = M - \bigcup_s B_{c(t+s)}(s)$ is called the half space associated with $c : [t, \infty) \to M$ (cf. figure 4.3).

Let $c_0 : [0, \ell] \to M$ be a geodesic in M with $c_0(0) = q_1, c_0(\ell) = q_2 \in H_t(c)$. If $c_0(r) \in B_{c(t+s_0)}(s_0)$ for some s_0 then $\mathrm{dist}(c_0(r), c(t+s)) \leq s - \epsilon$ for some $\epsilon > 0$ and all $s \geq s_0$ by the triangle inequality. Since $\mathrm{dist}(p_i, c(t+s) \geq s$ for all $s \geq 0$ this contradicts Toponogov's theorem 2.1 (T). Thus $H_t(c)$ is totally convex for all $t \geq 0$.

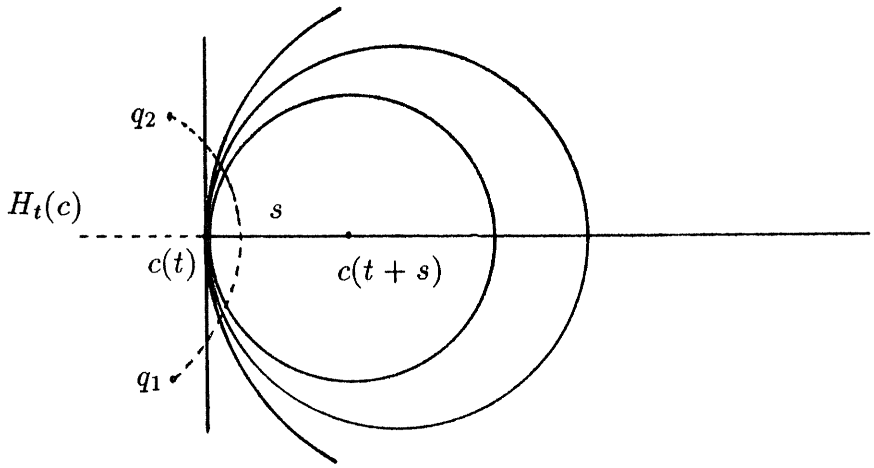


FIGURE 4.3

Now consider the compact totally convex subsets $C_t = \underset{c}{\cap} H_t(c)$, where the intersection is taken over all rays c emanating from p. Clearly $\cup_{t \geq 0} C_t = M$, and for each $t > 0$, $\partial C_t \neq \emptyset$.

If $C \subset M$ is a convex set with $\partial C \neq \emptyset$ then $\mathrm{dist}_{\partial C} : C \to \mathbb{R}$ is a concave function because sec $M \geq 0$, [10]. In particular $C_a = \{x \in C | \mathrm{dist}_{\partial C}(x) \geq a\}$ is convex for all $a \leq a_0 = \max \mathrm{dist}_{\partial C}$. Moreover dim $C^{a_0} <$ dim C. This construction terminates in finitely many steps when one arrives at a convex set $S \subset C$ without boundary.

By applying the above construction to one of the sets C_t, $t > 0$ we get a totally geodesic, closed submanifold $S \subset C_t \subset M$. Such an S is called a soul of M. It is now easy to see that each $x \in M - S$ lies on the boundary of some convex set $C \supset S$. By the convexity principle 1.3, x is a regular point for $\mathrm{dist}_S : M \to \mathbb{R}$. The soul theorem is then a consequence of 1.9 and 1.10. □

It is not known if the converse to 4.2 holds, i.e.

Problem 4.4. *Does the total space E of any vector bundle $E \to M$ over a closed riemannian manifold M with* sec $M \geq 0$, *carry a complete metric with* sec $E \geq 0$?

This problem remains unsolved even when $M = S^n$ (cf. [10], [54], [55]).

In trying to extend some of these ideas to complete noncompact manifolds M with sec $M \geq k$, $k < 0$, we reinterpret some of the above constructions for $k = 0$.

Specifically we observe that $\mathrm{dist}_p : M \to \mathbb{R}$ has no critical points outside the compact set $C_t = \bigcap_c H_t(c)$ for any $t \geq 0$. In fact any point $x \in M - C_t$ belongs to $\bigcup_s B_{c(t+s)}(s)$ for some ray $c : [0, \infty) \to M$ with $c(0) = p$. Now $\bigcup_s B_{c(t+s)}(s) \subset \bigcup_s \mathrm{reg}(c(t), c(t+2s))$ by 2.1 and the regularity principle 2.5 applies.

With this in mind suppose M is a complete non compact riemannian manifold with $\sec M \geq -k, k > 0$ and finitely many ends. For fixed $p \in M$ and $r > 0$ let $R(p,r) = \{c(r) | c : [0, \infty) \to M \text{ ray}, c(0) = p\} \subset \mathrm{dist}_p^{-1}(r)$. Following [7] we define the essential diameter of ends at distance r from p, $\mathrm{esdi}(p,r)$ as

$$\mathrm{esdi}(p,r) = \sup \mathrm{diam} \sum_r \tag{4.5}$$

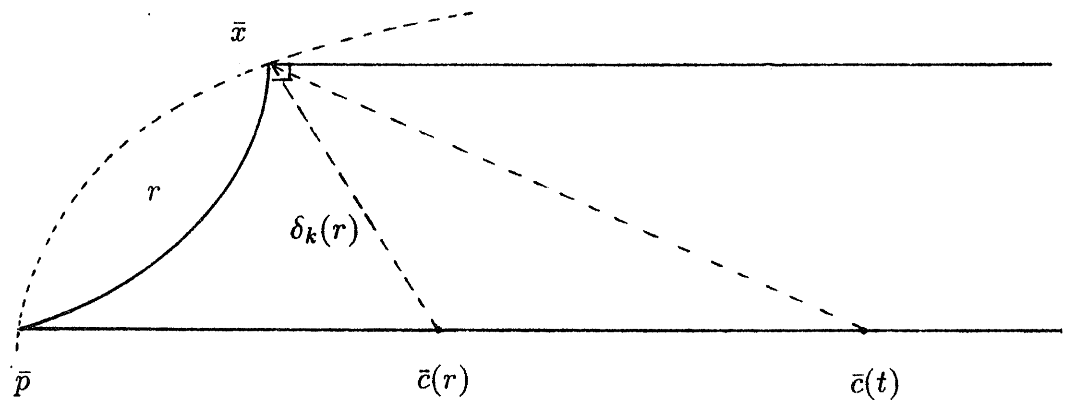
where the supremum is taken over all connected components $\sum_r$ of $\mathrm{dist}_p^{-1}(r)$ with $\sum_r \cap R(p,r) \neq \emptyset$ (corresponding to the boundary of unbounded components of $M - D(p,r)$). The *essential end diameter* of M is then

$$\mathrm{esdi}_\infty(M) = \inf_p \limsup_{r \to \infty} \mathrm{esdi}(p,r). \tag{4.6}$$

Of course $\mathrm{esdi}_\infty(M) = \infty$ in general. In view of the discussion of $\sec M \geq 0$ above, however, the following result of Shen [59] can be interpreted as a generalization of 4.1.

Theorem 4.7 (Bounded End Theorem). *Suppose M is a complete non compact riemannian manifold with finitely many ends and $\sec M \geq -k, k > 0$. If $\mathrm{esdi}_\infty(M) < \frac{1}{\sqrt{k}} \log(\frac{3+\sqrt{5}}{2})$, then M is diffeomorphic to the interior of a compact manifold with boundary.*

Proof. By assumption there is a $p \in M$ and an $R > 0$ so that $\mathrm{diam} \sum_r < \frac{1}{\sqrt{k}} \log(\frac{3+\sqrt{5}}{2})$ for all connected components $\sum_r \subset \mathrm{dist}_p^{-1}(r)$ with $\sum_r \cap R(p,r) \neq \emptyset, r \geq R$. In particular $\mathrm{dist}(x, R(p,r)) < \frac{1}{\sqrt{k}} \log(\frac{3+\sqrt{5}}{2})$ for all x in such $\sum$'s. An easy application of the regularity principle 2.5 (cf. figure 4.8), then shows that all points of $\sum_r \subset \mathrm{dist}_p^{-1}(r)$ with $\sum_r \cap R(p,r) \neq \emptyset, r \geq R$ are regular points for p. If now U is one of the finitely many unbounded connected components of $M - D_p(R)$, there is a ray $c : [0, \infty) \to M$ emanating from p with $c(r) \in U$ for any $r > R$. The isotopy lemma 1.8 then implies that all $\sum_r, r \geq R$ determined by $c(r) \in \sum_r$ are homeomorphic and in fact U is homeomorphic to $\sum_R \times (R, \infty)$. From this and the fact that $M - D_p(R)$ has at most finitely many bounded components, we conclude that there is an $R_1 \geq R$ so that all points in $M - B_p(R_1)$ are regular points for dist_p. □



In S^2_k, $\mathrm{reg}(\bar{p}, \bar{c}) = \bigcup_t \mathrm{reg}(\bar{p}, \bar{c}(t)) = \bigcup_r \mathrm{dist}_{\bar{p}}^{-1}(r) \cap B_{c(r)}(\delta_k(r))$

where $\delta_k(r) = \dfrac{1}{\sqrt{k}} \cosh^{-1}\left(\dfrac{\cosh^3 \sqrt{k}r - \sinh^3 \sqrt{k}r}{\cosh \sqrt{k}r}\right)$.

FIGURE 4.8

When the essential end diameter (4.6) is indefinite (or violates the assumption in 4.7), the regularity argument used above does not apply directly. However, existence of p-critical points in $M - B_p(r)$ for any r, restricts suitable excess-invariants of M severely. This was first observed in [3] (cf. also [59] and [7]. Abresh and Gromoll also made the fundamental discovery that excess functions can be controlled in terms of Ricci curvature. If the end diameter growth function 4.5 is not too big, this excess control, then violates the restrictions stemming from the existence of critical points in $M - B_p(r)$ for all r. More precisely [3],

Theorem 4.9. *Let M be a complete non compact riemannian n-manifold with* sec $M \geq -k, k > 0$ *and end diameter growth satisfying*

$$\inf_p \limsup_{r\to\infty} esdi(p,r) \cdot r^{-\frac{1}{n}} < \frac{1}{8} k^{-\frac{n-1}{2n}}.$$

Then M is diffeomorphic to the interior of a compact manifold with boundary if also $Ric M \geq 0$.

For other similar results involving intermediate curvatures and volume restrictions see [59] and [62].

It should also be mentioned here, that examples due to Sha and Yang [58], show that the diameter growth assumption in 4.9 is necessary. For more details about all of this we refer to the survey article by Gromoll [19] in these proceedings.

Remark 4.10. In the framework of this survey, there are so far no general structure theorems for natural classes of closed manifolds, e.g. positively or non negatively curved manifold. There is reasonable hope, however, that results of this type may emerge via our increased understanding of the Gromov-Hausdorff topology on such natural classes of closed manifolds. For more about this we refer to the surveys by Petersen [52], and Fukaya [13].

5. FINITENESS THEOREMS

It is not unreasonable to view a finiteness theorem as a first step in a recognition process, or for that matter as an approximation to a structure theorem.

In the results we will discuss here, the following simple packing lemma plays a crucial role.

Lemma 5.1 (Packing Lemma). *Fix an integer $n \geq 2$, and $k \in \mathbb{R}$. For every $\epsilon, D > 0$ there is an $N = N(n,k,\epsilon,D)$ such that any D-ball in a riemannian n-manifold M with $Ric\, M \geq k(n-1)$ can be covered by $\leq N$ ϵ-balls.*

Proof. Let $\{x_i, \dots, x_N\}$ be the centers of a maximal set of disjoint $\epsilon/2$-balls in the given D-ball in M. By maximality, the corresponding ϵ-balls provide a cover.

Moreover, the relative volume comparison lemma 2.4 yields

$$N \leq v_k^n(D)/v_k^n(\epsilon/2) = N(n,k,\epsilon,D).$$

□

It is useful to know that when an ϵ-cover is chosen as efficiently as in the proof of 5.1, the number of ϵ-balls with non-empty intersection is also a priori bounded independent of ϵ ! This is another simple consequence of 2.4 (cf. e.g. [31]).

We now have all the ingredients used in the proof of the following result (cf. [31]).

Theorem 5.2 (Homotopy Finiteness). *Given an integer $n \geq 2$, a real k and positive D, v. There are at most finitely many homotopy types among closed riemannian n-manifolds M satisfying* $\sec M \geq k$, *$\operatorname{diam} M \leq D$ and $\operatorname{vol} M \geq v$.*

Proof. By the local geometric contractibility lemma 3.11 (3.12), there are constants $r = r(n,k,D,v) > 0$ and $L = (n,k,D,v) > 0$, such that any pair $(p_1,p_2) \in M \times M$ with $\operatorname{dist}(p_1,p_2) = \epsilon \leq r$ is joined by a path $\sigma_{p_1p_2} : [0,1] \to M$ depending continuously on (p_1,p_2) and with $\operatorname{Length}(\sigma_{p_1p_2}) \leq L \cdot \epsilon$. This allows us to think of $\sigma_{p_1p_2}(t)$ as the "center of mass" of the points p_1, p_2 with "weights" $1-t$ and t respectively. For any integer $\ell \geq 2$ it is clear that by iterating this construction, we can assign to any ordered ℓ-tuple $(p_1,\ldots,p_\ell)$ in M a map $\sigma_{p_1,\ldots,p_\ell} : \Delta_\ell \to M$ of the standard $\ell -$ simplex $\Delta\ell$, "spanning" $(p_1,\ldots,p_\ell)$ and depending continuously on $(p_1,\ldots,p_\ell)$, whenever $\operatorname{dist}(p_i,p_j) \leq \epsilon = \epsilon(\ell,r,L) \leq r$, $i,j = 1,\ldots,\ell$.

As pointed out after 5.1 there is an $\ell = \ell(n,k,D)$ such that in any "efficient" cover of M by ϵ-balls at most ℓ such balls meet, independent of ϵ. With this ℓ, choose $\epsilon = \epsilon(\ell,r,L) = \epsilon(n,k,D,v)$ as above and fix for each closed riemannian n-manifold M with $\sec M \geq k, \operatorname{diam} M \leq D$ and $\operatorname{vol} M \geq v$, an "efficient" cover by $\epsilon/2$-balls, $B(p_1,\epsilon/2),\ldots,B(p_m,\epsilon/2)$. According to the packing lemma 5.1, $m \leq N = N(n,k,D)$.

For each $m \in \{1, \ldots, N\}$ consider all the manifolds M that are covered by exactly m $\epsilon/2$-balls as above, and where the corresponding nerves have isomorphic 1-skeleta, i.e. ball number i intersects ball number j in all M, or ball number i does not intersect ball number j in any M. This obviously divides our class of manifolds M into finitely many subclasses. From the above choices and constructions it now follows that by choosing a partition of unity subordinate to each fixed $\epsilon/2$-cover we get continuous maps, f, f' between any two manifolds M, M' from the same subclass. Moreover, by construction $\operatorname{dist}(f' \circ f, id_M) \leq C \cdot \epsilon$, for some $C = C(L, \ell) = C(n, k, D, v)$. Making sure that also $C \cdot \epsilon \leq r$, then shows that our maps f, f' are homotopy equivalences. This completes the proof. □

The conclusion in 5.2 can actually be sharpened considerably (cf. [36] and the announcement in [6] of recent work of Perelman). This relies heavily on Gromov-Hausdorff convergence techniques and geometric topology, and goes beyond the main topic discussed here (see [52]).

Theorem 5.3 (Topological Finiteness). *For fixed n, k, D and v as above the class of closed riemannian n-manifolds M with $\sec M \geq k, \operatorname{diam} M \leq D$ and $\operatorname{vol} M \geq v$ contains at most finitely many homeomorphic types.*

Since any closed topological manifold of dimension $n \neq 4$ has at most finitely many differentiable structures [44], the above theorem yields actually *finiteness of diffeomorphism types* when $n \neq 4$. Therefore 5.3 generalizes Cheeger's finiteness theorem [8] except for $n = 4$.

Question 5.4. *Are there $k \in \mathbb{R}, D, v > 0$, and infinitely many diffeomorphism types of closed riemannian 4-manifolds M satisfying $\sec M \geq k, \operatorname{diam} M \leq D$ and $\operatorname{vol} M \geq v$?*

The volume assumption in 5.2 (5.3) prevent the phenomenon called collapsing. It is remarkable that even without this assmuption, one has the following general finiteness theorem due to Gromov [23].

Theorem 5.4 (Betti Number Theorem). *Given an integer* $n \geq 2$, *a real* k *and positive* D. *There is a* $C = C(n, k, D)$ *so that for every field,* F *of coefficients,* $\dim H_*(M; F) \leq C$ *when* M *is any closed riemannian* n- *manifold satisfying* $\sec M \geq k$ *and* $\operatorname{diam} M \leq D$.

Note that since $\sec M \geq 0$ is a scalings invariant property, it follows that in this theorem the diameter assumption is superfluous when $k = 0$.

Besides the original proof of this theorem in [23] (cf. also [1],[2]), detailed proofs have been given in lecture notes by Cheeger [7] and Meyer [46]. Here, therefore, we will only attempt to elaborate on the main ideas and strategies of the proof:

First observe that if for some reason one could control the homology of all ϵ-balls for some a priori ϵ, then the packing lemma combined with a Mayer-Victoris type argument would give control on the homology of all M^n with $\sec M \geq k$ and $\operatorname{diam} M \leq D$. For fixed k, however, there is a small $\epsilon = \epsilon(k)$ so that all comparison arguments applied to such ϵ-balls are essentially the same as comparison arguments used for $k = 0$. This then "reduces" the proof to the case $\sec M \geq 0$ and no diameter assumption. Now, however, one needs to be able to control the homology of all balls B. It is important that all of the above ideas work if one replaces $\dim H_*(B)$ (which could be infinite) by the finite number

$$\operatorname{cont} B = \operatorname{rank}(i_* : H_*(B) \to H_*(5B)),$$

called the *content* of B. Here $5B$ is concentric with B and has 5 times larger radius. The only significance of 5 is that it is a fixed number > 1.

In order to estimate $\operatorname{cont}(B)$, one makes the following simple but basic observation: If B can be isotoped inside $5B$ to a ball B' with $5B' \subset 5B$, then

$$\operatorname{cont} B \leq \operatorname{cont} B'.$$

In this case we say that B can be *compressed* to B'. If $\operatorname{rad} B' \leq \frac{1}{2} \operatorname{rad} B$ we say that B is *compressible*. Thus, B is *incompressible* if it does not compress to any B' with $\operatorname{rad} B' \leq \frac{1}{2} \operatorname{rad} B$.

Obviously, if a ball B can be compressed sufficiently many times, so as to end up in a contractible ball (e.g., a ball with radius smaller than $\mathrm{inj}(M)$) its content is 1.

If on the other hand $B_p(r)$ is incompressible then by the isotopy lemma 1.8 any point $x \in 2B$ has a critical value in the interval $[\frac{1}{2}r, r + \mathrm{dist}(x,p)]$ (cf. figure 5.5).

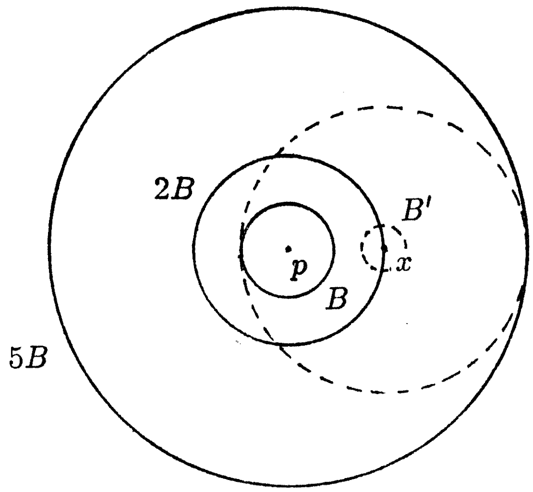


FIGURE 5.5

Now cover $B_p(r)$ efficiently with balls of radius say $r/10$ as in the packing lemma 5.1. If all of these balls have content 1 we get an a priori estimate for $\mathrm{cont}\, B_p(r)$. Otherwise we find a $q \in B_p(\frac{11}{10}r)$ so that all $x' \in 2B_q(\frac{1}{10}r'), r' \le r$ have critical values in the interval $[\frac{1}{20}r', \frac{1}{10}r' + \mathrm{dist}(x', q)]$ as well as in the interval $[\frac{1}{2}r, r + d(x', p)]$ from before. By the criticality principle 2.8, this process must terminate after a finite a priori number of steps where all involved balls have content 1. The a priori control on the number of such balls needed, then gives via the Mayer-Victoris type argument alluded to above the desired conclusion. □

We point out that the constructive proofs discussed above yield explicit estimates, although probably very far from being sharp. The following conjecture has been formulated by Gromov.

Conjecture 5.6. *Any complete riemannian n-manifold M with* $\sec M \ge 0$ *must have dim* $H_*(M; F) \le \dim H_*(T^n; F) = 2^n$.

Another natural question is whether the product structure of $H_*(M; F)$ is controlled for manifolds M in the class sec $M \geq k$ and diam $M \leq D$. In particular

Question 5.7. *Given an integer $n \geq 2, k \in \mathbb{R}$ and $D > 0$. Are there only finitely many rational homotopy types among closed simply connected riemannian n-manifold M, satisfying* sec $M \geq k$ *and* diam $M \leq D$?

REFERENCES

1. U. Abresch, *Lower curvature bounds, Toponogov's theorem, and bounded topology, I*, Ann. Sci. Ecole Norm. Sup. **18** (1985), 563–633.
2. ______, *Lower curvature bounds, Toponogov's theorem, and bounded topology, II*, Ann. Sci. Ecole Norm. Sup. **20** (1987), 475–502.
3. U. Abresch and D. Gromoll, *On complete manifolds with non-negative Ricci curvature*, J. Amer. Math. Soc. **3** (1990), 355–374.
4. A. L. Besse, *Manifolds all of whose geodesics are closed*, Ergeb. Math. Grenzgeb. **93** (1978), Springer-Verlag.
5. J. Bernstein and T. Ganea, *Homotopical nilpotency*, Illinois J. Math. **5** (1961), 99–130.
6. Y. Burago, M. Gromov and G. Perelman, *A. D. Aleksandrov's spaces with curvature bounded from below, I*, preprint.
7. J. Cheeger, *Critical points of distance functions and applications to geometry*, Springer Lecture notes, to appear.
8. ______, *Finiteness theorems for Riemannian manifolds*, Amer. J. Math. **92** (1970), 61–74.
9. J. Cheeger and D. G. Ebin, *Comparison Theorems in Riemannian Geometry*, North Holland Math Library **9**, Amsterdam, 1975.
10. J. Cheeger and D. Gromoll, *On the structure of complete manifolds of non-negative curvature*, Ann. of Math. **96** (1972), 413–443.
11. O. Durumeric, *Manifolds of almost half of the maximum volume*, Proc. Amer. Math. Soc. **104** (1988), 277–283.
12. J.-H. Eschenburg, *Diameter, volume and topology for positive Ricci curvature*, J. Differential Geometry **33** (1991), 743–747.
13. K. Fukaya, *Hausdorff convergence of Riemannian manifolds and its applications*, preprint.
14. R. Greene and P. Petersen V, *Little Topology, Big volume*, preprint.
15. R. Greene and K. Shiohama, *Convex functions on complete non compact manifolds: topological structure*, Invent. Math. **63** (1981), 129–157.
16. ______, *Convex functions on complete non compact manifolds: differentiable structure*, Ann. Sci. Ecole Norm. Sup. **14**, 357–367.
17. R. Greene and H. Wu, *C^∞ convex function and manifolds of positive curvature*, Acta Math. **137** (1976), 209–245.
18. ______, *Integrals of subharmonic functions on manifolds of nonnegative curvature*, Invent. Math. **27** (1974), 265-298.
19. D. Gromoll, *Spaces of nonnegative curvature*, these proceedings.
20. D. Gromoll, W. T. Meyer, and W. Klingenberg, *Riemannsche Geometrie im Grossen*, Lecture Notes in Math., vol. 55, Springer-Verlag, 1968.
21. D. Gromoll and K. Grove, *A generalization of Berger's rigidity theorem for positively curved manifolds*, Ann. Sci. Ecole Norm. Sup. **20** (1987), 227–239.
22. ______, *The low-dimensional metric foliations of euclidean spheres*, J. Differential Geometry **28** (1988), 143–156.
23. M. Gromov, *Curvature, Diameter and Betti numbers*, Comment Math. Helv. **56** (1981), 179–195.
24. ______, *Filling Riemannian Manifolds*, J. Differential Geometry **18** (1983), 1–148.
25. ______, *Manifolds of negative curvature*, J. Differential Geometry **13** (1978), 223–230.