Comparison Geometry
MSRI Publications
Volume **30**, 1997

A Genealogy of Noncompact Manifolds of Nonnegative Curvature: History and Logic

ROBERT E. GREENE

ABSTRACT. This article presents an approach to the theory of open manifolds of nonnegative sectional curvature via the calculus of nonsmooth functions. This analytical approach makes possible a very compact development of the by now classical theory. The article also gives a summary of the historical development of the subject of open manifolds of nonnegative sectional curvature and of related topics. Some very recent results are also discussed, including results of the author jointly with P. Petersen and S. H. Zhu on curvature decay.

Introduction

At first sight, the study of noncompact manifolds seems necessarily more complicated than that of compact ones: Removal of a single point from a compact manifold gives a noncompact one, but not all noncompact manifolds arise in this way—in general, the topological one-point compactification of a noncompact C^∞ manifold is not even a topological manifold. To look at the matter another way, more relevant to the subject of this article, a C^∞ compact manifold always admits a C^∞ function with only nondegenerate critical points, and compactness implies that there are only a finite number of such critical points; thus, a compact manifold has the homotopy type of a finite CW-complex, or what we shall call loosely finite topology. A C^∞ noncompact manifold (all manifolds will be C^∞ from now on) of course also always admits a proper C^∞ function with only nondegenerate critical points, but for some manifolds all such proper functions have infinitely many critical points. (Here *proper* means that the inverse of each set of the form $(-\infty, \alpha]$ is compact: this is the natural context for Morse theory.) Correspondingly, the manifold may fail to have finite topology, that is, it may not have the homotopy type of a finite complex.

Research supported by the National Science Foundation.

These elementary observations can be given more considerable depth. There is, for instance, a sense in which even just noncompact Riemann surfaces, let alone Riemannian manifolds of higher dimension, cannot be described by invariants, without special set-theoretic assumptions [Becker et al. 1980].

All this changes, however, if attention is restricted to manifolds admitting a complete Riemannian metric of nonnegative curvature.

Completeness itself imposes no restrictions: Every (paracompact) manifold admits a complete Riemannian metric, as can be seen by elementary arguments. Completeness alone does, however, yield a geometrically interesting object, as follows. If M is a complete, noncompact Riemannian manifold and $x_0 \in M$ is any point, we can take a sequence $\{x_i\}$ in M such that $\lim_{i\to+\infty} \operatorname{dist}(x_0, x_i) = +\infty$; if no such sequence existed, M would be bounded and hence compact. Now consider geodesics γ_i, parametrized by arclength, from x_0 to x_i, with $\gamma(0) = x_i$ and $\gamma(\operatorname{dist}(x_0, x_i)) = x_i$, and take a subsequence $\{\gamma_{i_j}\}$ such that $\{\gamma'_{i_j}(0)\}$ converges, say to V (such a subsequence must exist by the compactness of the unit ball in $T_{x_0}M$). Then the geodesic defined on $[0, +\infty)$ by

$$\gamma : t \mapsto \exp_{x_0}(tV)$$

is what is called a *ray*, that is, an arclength parameter geodesic defined on $[0, +\infty)$ that is a minimal-length connection between any two of its points. In Mark Twain's phrase, γ "lights out for the Territory" as fast as possible: intuitively, γ is a shortest connection to infinity.

In the absence of curvature hypotheses, the existence of a ray is no more than a geometric version of the essentially obvious fact that there is always a proper, injective map of $[0, +\infty)$ into a given noncompact manifold. But in the presence of everywhere nonnegative curvature, a new kind of meaning arises. Nonnegativity, and especially positivity, of curvature tends to make long geodesics nonminimizing. Thus some tension arises between the necessary existence of rays and the curvature's nonnegativity. This tension ultimately imposes strong restrictions on the topology of the manifolds that admit complete metrics of nonnegative curvature.

Completeness is essential here: every noncompact manifold admits a (generally noncomplete) metric of nonnegative sectional curvature [Gromov 1986].

Historically, the precise forms of these matters were discovered rather slowly and, as it were, almost indirectly in some cases. The historical development is shown on the genealogical chart on pages 132–134, albeit the lines of descent are from the imagination of the present author to some extent. The theme of the present article is, however, that, with the benefit of hindsight, and with the use of certain principles of analysis—which indeed could have been used from the start, very nearly—the whole subject can be developed very rapidly, essentially as a repeated application of the second variation formula. Thus, there is a certain contrast, though by no means a dichotomy, between the genealogy of the historical development shown in the chart, and the intrinsic logic, as the author perceives it at least, of the subject independent of history.

1. Of Second Variation and How Positive Curvature Can Be

Euclidean space with its standard flat metric is the separating case between spherical (constant positive sectional curvature) and hyperbolic (constant negative) geometries. This is a familiar classical concept, but it has certain subtle aspects. In particular, straight lines, which can be thought of as limits of longer and longer geodesics in larger and larger spheres, with curvature going to zero, are, from this viewpoint, just barely minimizing. One can attach precise meaning to this in the following way:

Consider a geodesic with arclength parameter, $\gamma : [0, L] \to M$, in some Riemannian manifold. And consider variations of γ with the variation vector field V along γ being everywhere perpendicular to γ (or equivalently to γ'). Suppose the variation has fixed endpoints so that $V(0) = 0$ and $V(L) = 0$. Finally, as a normalization, suppose that $\max_{t \in [0,L]} \|V(t)\| = 1$. Within this class of variations, the infimum of the second variation is in effect a measure of to what extent γ is minimizing.

Now what is interesting from the present viewpoint is that this infimum, for straight lines in Euclidean space, goes to 0 as L goes to infinity. (Its exact value is $4/L$.) The limit is nonzero (and positive) for a hyperbolic space, whereas for a sphere, negative second variation is of course possible for sufficiently long geodesics, that is sufficiently large values of L, implying, as is true, that such long geodesics are not of minimum length among nearby curves with the same endpoints.

The fact that straight lines in $\mathbb{R}^n$ are, in this sense, just barely minimizing suggests that there should be quantitative estimates on just how much positivity of curvature could be possible without forcing a geodesic to be nonminimal. The following lemma gives a specific version of this:

LEMMA 1.1. *Let $\gamma : [0, L] \to M$ be an arclength parameter geodesic and $N(t)$ be a parallel unit vector field along γ with $\langle N(t), \gamma'(t) \rangle \equiv 0$ for all $t \in [0, L]$. Let $k(t)$ be the sectional curvature of the two-plane spanned by $N(t)$ and $\gamma'(t)$. If $a, b > 0$ and $a + b = L$ and if γ is minimizing among nearby curves connecting its endpoints*, *then*

$$\frac{1}{a^2} \int_0^a t^2 k(t)\, dt + \frac{1}{b^2} \int_0^b t^2 k(L-t)\, dt \leq \frac{1}{a} + \frac{1}{b}.$$

PROOF. Define a piecewise C^∞ vector field $V(t)$ along γ by

$$V(t) = \begin{cases} (t/a)N(t) & \text{if } 0 \leq t \leq a, \\ ((L-t)/b)N(t) & \text{if } a \leq t \leq L. \end{cases}$$

Then define a variation of γ by

$$(t,s) \mapsto \exp_{\gamma(t)}(sV(t)).$$

Since γ is locally minimizing, the associated second variation of arclength must be nonnegative. By the usual formula, this becomes

$$\begin{aligned} 0 &\le \int_0^L \langle D_t V, D_t V\rangle\, dt - \int_0^L \|v\|^2 k(t)\, dt \\ &= a\,\frac{1}{a^2} + b\,\frac{1}{b^2} - \frac{1}{a^2}\int_0^a t^2 k(t)\, dt - \frac{1}{b^2}\int_a^L (t-a)^2 k(t)\, dt \\ &= \frac{1}{a} + \frac{1}{b} - \frac{1}{a^2}\int_0^a t^2 k(t)\, dt - \frac{1}{b^2}\int_0^b t^2 k(L-t)\, dt. \qquad \square \end{aligned}$$

This lemma has two important corollaries. The first is an aspect of the Toponogov Splitting Theorem. The second, which is in effect a restriction on how much positivity of curvature there can be, is not so familiar, but forms an important complement to the results in Section 6 on how little positivity of curvature there can be.

COROLLARY 1.2 (TOPONOGOV). *Let $\gamma : (-\infty, +\infty) \to M$ be an arclength parameter geodesic with* $\operatorname{dist}(\gamma(t_1), \gamma(t_2)) = |t_1 - t_2|$ *for all $t_1, t_2 \in (-\infty, +\infty)$. If, for some parallel unit vector field along γ, we have $k(t) \ge 0$ for all t, then $k(t) \equiv 0$ for all t.*

This follows immediately from the lemma by choosing $\gamma(a)$ of the lemma to correspond to $\gamma(t)$ of the corollary, for a given fixed t, and then letting a and b go to infinity.

By choosing $n-1$ parallel, mutually perpendicular unit vector fields along γ, each perpendicular to γ', one easily obtains results corresponding to the lemma and corollary for Ricci curvature in place of sectional curvature.

For the second corollary, the following notation is needed: For a fixed $p \in M$, let $K_p(t)$ be the infimum of the sectional curvatures occurring at points q with $\operatorname{dist}(q,p) = t$.

COROLLARY 1.3. *If M is a noncompact complete manifold of nonnegative sectional curvature and $p \in M$, then for each $r > 0$ we have*

$$\frac{1}{r}\int_0^r t^2 K_p(t)\, dt \le 1.$$

This follows from Lemma 1.1 by choosing a ray $\gamma : [0, +\infty) \to M$ with $\gamma(0) = p$, setting $a = r$ and letting b go to infinity.

In terms of uniformly behaved $K_p(t)$, Corollary 1.3 makes clear that quadratic decay of $K_p(t)$ as a function of t is the critical case: $K_p(t)$ being of order $1/t^2$ gives the exact order of magnitude that the corollary allows.

This quadratic decay does in fact occur: For the paraboloid of revolution obtained by revolving the curve $y = \sqrt{x}$ around the x-axis, K_p equals $1/(2x+\frac{1}{2})^2$.

Here x and the arclength t from the origin $p = (0, 0)$ are uniformly comparable (since the derivative of $\sqrt{x}$ is bounded as $x \to +\infty$, and indeed goes to 0), so that $K(t)$ also decays quadratically as a function of t.

It is important to understand, however, that these observations about decay rates are really only relevant to cases where $K(t)$ has quite regular behavior. By way of example, note that even when M is a rotationally symmetric surface, with a metric in geodesic polar coordinates of the form $dr^2 + f^2(r)\, d\theta^2$, it is still possible for $\limsup_{r\to+\infty} K_p(r)$ to be $+\infty$ (and curvature to be everywhere > 0). An example can be constructed as a surface of revolution.

2. Of Analysis, Support Functions, and Comparisons

Curvature provides estimates on second variations of arclength, which one naturally thinks of as estimates (albeit in general one-sided only) on the second derivatives of distances. But this natural viewpoint does not apply literally, without some explication, since distances are not in general smooth functions on Riemannian manifolds. Fortunately, this lack of smoothness turns out not to matter, really. The essential reasoning needed to obviate the requirement of smoothness has deep roots from long ago in real and complex analysis, for instance, in the idea that subharmonicity needed to be defined and used for functions lacking any differentiability. Indeed, so widespread are this and other similar concepts in analysis that it is perhaps surprising that they began to appear explicitly in Riemannian geometry, it seems, only as recently as in [Calabi 1957]. Since then, however, they have been systematically studied and exploited in both Riemannian and Kähler contexts, in particular, in [Cheeger and Gromoll 1971; 1972; Greene and Wu 1976; 1978; Elencwajg 1975; Wu 1970; 1987].

The basic viewpoint needed is simply this: Define a function g to be *supporting* a function f at a point p if $f(p) = g(p)$ and $g(q) \le f(q)$ for all q in a neighborhood of p. Then:

SUPPORT PRINCIPLE FOR CONVEXITY. *A continuous function f on a Riemannian manifold is convex* (*that is*, *convex along each geodesic*) *if it is supported at each point p by a convex function on some neighborhood of p.*

The proof is elementary, since in effect one can immediately reduce it to a standard and easy result about convex functions on the real line. In practice, a somewhat refined principle is more useful.

Define a C^∞ function f to be *ε-convex* if the eigenvalues of its second covariant differential are everywhere at least $-\varepsilon$. (Recall that the second covariant differential is the quadratic form $D^2_f(X, Y) = X(Yf) - (D_X Y)f$.) Then:

REFINED SUPPORT PRINCIPLE FOR CONVEXITY. *A continuous function f on a Riemannian manifold is convex if*, *for each point p and each $\varepsilon > 0$*, *there is an ε-convex function g defined on some neighborhood of p such that g supports f at p.*

The refined principle follows from the original principle easily using the fact that locally defined C^∞ strongly convex functions always exist on Riemannian manifolds; for example, $\operatorname{dist}^2(x_0, \cdot\,)$ is C^∞ in a neighborhood of x_0 and the second covariant differential of this function has eigenvalues equal to 2 at x_0, and hence bounded away from 0 in a neighborhood U of x_0. The original principle can be applied to show that, for fixed but arbitrary x_0, and every $\delta > 0$, the functions $f(\cdot\,)+\delta\operatorname{dist}^2(x_0, \cdot\,)$ are convex on U under the hypotheses of the refined principle. Since convexity is a local property, and since a limit of convex functions is convex, the conclusion of the refined principle follows by letting $\delta \to 0^+$.

The support principle can also be established for subharmonic functions, and also for plurisubharmonic functions on complex manifolds. Only the subharmonic case will be discussed here: for the complex case, see [Elencwajg 1975] and [Greene and Wu 1978].

By definition, a (real-valued) continuous function f on a Riemann manifold is *subharmonic* if it is a subsolution of Dirichlet problems, that is, if $f(x) \le h(x)$ for all $x \in K$, where K is any compact set with nonempty interior and $h : K \to \mathbb{R}$ is any continuous function harmonic in $\mathring{K}$ and satisfying $h(q) \ge f(q)$ for $q \in K - \mathring{K}$. We define a C^∞ function g to be *ε-subharmonic* if $\Delta g \ge -\varepsilon$ wherever g is defined (here $\Delta = \sum \partial^2/\partial x_i^2$ at the center of a normal coordinate system $x_1, \dots, x_n$). Then:

Refined support principle for subharmonicity. *A continuous function f on a Riemannian manifold is subharmonic if, for each $p \in M$ and $\varepsilon > 0$, there is an ε-subharmonic function g defined on some neighborhood U of p such that g supports f at p.*

The proof runs almost parallel to that of the refined principle for convexity. On some (smaller) neighborhood V of p, the function $\operatorname{dist}^2(p, \cdot\,)$ has positive Laplacian, bounded away from 0. If h is a (C^∞) function that is harmonic in a neighborhood of p, then for each $\delta > 0$, $f + \delta\operatorname{dist}^2(p, \cdot\,) - h$ is supported at each point in a fixed neighborhood of p by the sum of a subharmonic function and a C^∞ function with positive Laplacian; this follows by choosing $\varepsilon > 0$ sufficiently small and using the function $g + \delta\operatorname{dist}^2(p, \cdot\,) - h$, where g is as in the hypothesis. Thus, by calculus, $f + \delta\operatorname{dist}^2(p, \cdot\,) - h$ cannot have a local maximum near p. Hence $f + \delta\operatorname{dist}^2(p, \cdot\,)$ is subharmonic. Hence, since the limit of subharmonic functions is obviously subharmonic, f is subharmonic. There will be circumstances later when not only convexity or subharmonicity can be deduced, but also even smoothness.

Smoothness principle. *Let f be a function on a Riemannian manifold. If f and $-f$ are both convex, or both subharmonic, then f is C^∞ and harmonic.*

Actually, the convex situation is a special case of the subharmonic one, since convex functions are always subharmonic [Greene and Wu 1973]. But it is nice to notice that the convex case can be proved by completely elementary methods:

The function must be linear along all geodesics. To prove differentiability in a neighborhood of p, choose a geodesic γ_1 through p and a point x_1 near p but not on γ_1. Linearity shows that f is C^∞ on the local two-dimensional submanifold generated by geodesics from x_1 to (nearby) points of γ_1. Repeating this cone construction with x_2 near p but not on the two-dimensional submanifold gives a three dimensional local submanifold (containing p) on which f is differentiable. Further repetition eventually gives differentiability on a neighborhood of p. So f is differentiable and, since linear along geodesics, harmonic.

To treat the general subharmonic case, one proceeds as follows: For a given point p, choose a small closed ball B around p. By standard results in partial differential equations, there is a harmonic function h on $\mathring{B}$, with h continuous on B and $h \equiv f$ on $B - \mathring{B}$. By the subharmonicity of f and $-f$, $h \equiv f$ on $\mathring{B}$. Hence f is C^∞ (and harmonic) on $\mathring{B}$.

The support function method gives an almost immediate proof of the Toponogov Comparison Theorem, for the case of comparison with euclidean space, as pointed out originally by H. Karcher [1989]. (The negative and positive curvature comparisons are also obtained in the same work by similar methods, but some small technical subtleties are involved, and these cases, which are not directly relevant here, will be omitted.) The crucial observation is the following lemma, which follows from the support principle and an easy second variation argument.

LEMMA 2.1. *If M is a complete Riemannian manifold with nonnegative sectional curvature, with p a point of M, and if $\gamma(t)$ is an arclength parameter geodesic in M, then the function $F : t \mapsto t^2 - \operatorname{dist}^2(p, \gamma(t))$ is convex.*

PROOF. It suffices to construct a C^∞ support function g_{t_0} for F at $t = t_0$, fixed but arbitrary, with $d^2 g_{t_0}/dt^2 \geq 0$. For then, given $\varepsilon > 0$, there is a neighborhood of t_0 on which g_{t_0} is ε-convex, and the refined support principle applies. To construct such a g_{t_0}, choose a minimal geodesic $\gamma_1 : [0, L] \to M$ from p to $\gamma(t_0)$, where $L = \operatorname{dist}(p, \gamma(t))$. Define a vector field $W(s)$ along $\gamma_1(s)$ as the parallel translation of $\gamma'(t_0)$ along γ_1. Set $V = (s/L)W$ and define a variation of γ_1 by

$$(s, t) \mapsto \exp_{\gamma_1(s)}(tV(s)).$$

Let $L(t)$ be the length of the curve, as s varies over $[0, L]$ and t is fixed. Then a standard second variation calculation shows that $d^2 L(t)^2/ds^2 \leq 2$. On the other hand, clearly $\operatorname{dist}^2(p, \gamma(t)) \leq (L(t))^2$. Thus $g(t) = t^2 - L(t)^2$ satisfies the required conditions. □

Toponogov's Comparison Theorem can be easily deduced from this lemma in the form that estimates the third side of a geodesic triangle in terms of the other two sides and the included angle. Specifically, one wants to prove the following:

THEOREM 2.2 (TOPONOGOV COMPARISON FOR $k \geq 0$). *Suppose $\gamma_1 : [0, L] \to M$ is a minimal geodesic from a point $p = \gamma_1(0)$ to a point $q = \gamma_1(L)$ in a complete*

Riemannian manifold M of nonnegative curvature, and γ is an arclength geodesic through q with $\gamma(0) = q$ and $\langle \gamma'(0), -\gamma_1'(L)\rangle = \cos\alpha$ (so α is the angle between $\gamma'(0)$ and $-\gamma_1'(L)$). Then, for each positive t,

$$\operatorname{dist}^2(p, \gamma(t)) \le t^2 + L^2 - 2Lt\cos\alpha.$$

Here the right-hand side of the inequality is of course the square of the length of the third side of a euclidean triangle whose other sides have lengths t and L and meet at an angle α.

PROOF. For each $\delta > 0$, the function

$$H_\delta(t) = t^2 + L^2 - 2Lt\cos\alpha + \delta t - \operatorname{dist}^2(p, \gamma(t))$$

is convex, by the lemma. Also $H(0) = 0$ and, by an easy first variation argument, $H(t)$ is positive for all small positive t. Thus $H_\delta(t)$ is positive for all positive t. Hence

$$H_0(t) = t^2 + L^2 - 2Lt\cos\alpha - \operatorname{dist}^2(p, \gamma(t)) \ge 0$$

for all positive t, which implies the conclusion of the theorem. □

3. Of Noncriticality and Nonnegative Curvature

The idea of support functions discussed in the previous section makes it possible to deal with what amounts to the second derivative properties of functions that do not have second derivatives in the literal sense. It is natural and often very useful to have in addition a way to deal with first derivative properties.

The functions that arise in geometry are almost all locally Lipschitz continuous, that is, for each compact set K there is some constant B_K such that, for all $x, y \in K$,

$$|f(x) - f(y)| \le B_K \operatorname{dist}_M(x, y).$$

Distance functions have this property by nature (and the triangle inequality), and geometrically constructed functions, which are derived in most cases from distance considerations, thus tend to have it, too. Lipschitz continuous functions on $\mathbb{R}$ are almost everywhere differentiable and equal the integral of their derivative, up to an integration constant. (These properties hold under the more general condition of absolute continuity, of course, but this additional generality is irrelevant for our present purposes.) Thus it is reasonable to try to relate general behavior of Lipschitz continuous functions to their first derivative properties.

For a function on a manifold, or what amounts to the same thing locally, a function of several variables, the most basic first derivative property is noncriticality, that is, nonvanishing of the gradient grad f, or, equivalently, of the differential df. This noncriticality has a natural possible generalization to the case of Lipschitz continuous functions. (This idea was introduced in [Greene and Shiohama 1981a]; see also [Greene and Wu 1974]. In the specific case of distance functions, related ideas occurred, for example, in [Grove and Shiohama 1977], but were tied to specifics of geodesic geometry.) In effect, one defines a point x to be a noncritical point of a function f if there is a continuously varying set of directions at and near x along which f decreases at a definite nonzero rate. (This set of directions would be $-\operatorname{grad} f$ in the case of C^1 functions f.)

To make this completely precise, define a vector $v \in T_xM$ to be a *subgradient* for a Lipschitz continuous function f if there is a continuous vector field V defined in a neighborhood of x with $V(x) = v$ and with the property that, for some $\varepsilon > 0$ and $\delta > 0$ and all y in a neighborhood of x,

$$f(\varphi_y(t)) - f(y) \leq -\varepsilon t$$

for all $t \in (0, \delta)$, where φ_y is the geodesic emanating from y with $\varphi_y'(0) = V(y)$. This property is independent of the choice of Riemannian metric; indeed, the only function of the geodesics φ_y is to provide a suitable continuously varying family of curves with specified initial tangent. The definition could also be made with integral curves of a local C^∞ vector field, but it is sometimes convenient to be able to deal with vector fields that are just continuous. All such variants of the definition yield equivalent concepts.

A point x is a *noncritical point* for a Lipschitz continuous function f if there is a subgradient for f in T_xM [Greene and Shiohama 1981b]. A point is *critical* for f if it is not noncritical. This definition of critical definitely does not coincide with the idea that f is "constant up to second order" at a critical point. For example, the function $x \mapsto \operatorname{dist}_M(x, p)$, for p fixed, always has p as a critical point, even though it increases at unit rate along arclength parameter geodesics emanating from p; similarly, if q is such that $\operatorname{dist}_M(p, q) = \sup_{x \in M} \operatorname{dist}_M(x, p)$, then q is a critical point even though there is a direction at q along which $\operatorname{dist}_M(\,\cdot\,, p)$ decreases at unit rate. The point is that such a direction of strict decrease (or increase) cannot be chosen to vary continuously in a neighborhood of q. Of course these definitions do coincide with the usual concept of being a noncritical point if f is C^1 differentiable.

Note that the property of being noncritical is open, almost by definition: If x is noncritical then so are all points y in a sufficiently small neighborhood of x, since the same V will work for such y as a subgradient. It is also easy to check, using the Lipschitz continuity of f, that if V is *any* continuous vector field defined in a neighborhood of x with $V(x)$ a subgradient at x for f, then V itself has the required property of the vector field in the definition of subgradient for some ε, δ and neighborhood of x. Finally, one can also check easily that a linear combination of subgradients with nonnegative coefficients and at least one positive coefficient is again a subgradient: this again uses the Lipschitz continuity of f.

The role of Lipschitz continuity in all this is primarily to insure that if two curves have nearly the same tangent vector at some common initial point, then the difference of f at parameter t along one curve from f along the other at parameter t is a small multiple of t. This provides the necessary control of one-sided (upper and lower) derivatives along curves.

Noncriticality in this extended sense is sufficient to make the main result of noncritical Morse theory work, to the extent that it could possibly work for nonsmooth functions:

LEMMA 3.1 (NONCRITICAL MORSE THEORY). *Suppose that* $f : M \to \mathbb{R}$ *is a Lipschitz continuous function.*

(1) *If, for some* $a, b \in f(M)$, *the set* $f^{-1}([a, b])$ *is compact and contains only noncritical points of* f, *then*
   (a) $f^{-1}((-\infty, a])$ *is homeomorphic to* $f^{-1}((-\infty, b])$, *and*
   (b) $f^{-1}((-\infty, a])$ *is homeomorphic to a* C^∞ *manifold-with-boundary* $(N, \partial N)$; *in particular* $f^{-1}(\{a\})$ *is homeomorphic to the* C^∞ *manifold* ∂N.

(2) *If, for some* $a > 0$ *in* $f(M)$, *the set* $f^{-1}([a, +\infty))$ *contains only noncritical points of* f *and* $f^{-1}([a, b])$ *is compact for all* $b > a$, *then* M *is diffeomorphic to* $f^{-1}((-\infty, a))$.

This lemma is established by smooth approximation techniques in [Greene and Shiohama 1981b]. The relevant smoothing ideas were introduced in [Greene and Wu 1973] and applied to noncritical point theory in [Greene and Shiohama 1981b]; compare [Greene and Wu 1974].

For literal distance functions $x \mapsto \operatorname{dist}_M(x, p)$ on a complete Riemannian manifold, the concept of noncriticality is equivalent to a condition on the geometry of geodesics: A point $x \neq p$ is noncritical if and only if there is a nonzero vector v in $T_x M$ such that, for every minimal geodesic $\gamma : [0, 1] \to M$ with $\gamma(0) = p$ and $\gamma'(1) = x$, the angle between $\gamma(1)$ and v is $> \pi/2$. It is an easy exercise in the first variation formula to see that this geometric geodesic condition is equivalent to the definition already given. Indeed, the geodesic condition was used as a definition in some early works, such as [Grove and Shiohama 1977].

Although analytically the concept of noncriticality seems almost primordial, it took some time for its utility in geometry to become widely appreciated. Early investigations include [Greene and Wu 1974; Grove and Shiohama 1977; Greene and Shiohama 1981a; 1981b]. A history and some further details are given in [Greene 1989].

The relevance of these concepts to the theory of manifolds of nonnegative curvature is twofold: First, it applies to distance functions themselves, at large distances. And second, there are nontrivial convex functions on noncompact manifolds of nonnegative curvature; and convex functions have no critical points other than the points of the global minimum set.

To treat the first, one notes the following.

LEMMA 3.2. *If* M *is a complete manifold of nonnegative sectional curvature, if* $p \in M$, *and if* $\delta > 0$, *then there is a* $B > 0$ *with the following property*: *if*

$\mathrm{dist}_M(p,x) \geq B$ and if γ is a minimal geodesic from p to x, with $\gamma(0)=p$, then there is a ray $\gamma_1 : [0,+\infty) \to M$ with $\gamma_1(0)=p$ and such that the angle between $\gamma'(0)$ and $\gamma_1'(0)$ is less than δ.

The proof is an easy modification of the basic limiting argument establishing the existence of rays. From this lemma and Toponogov's theorem, one can reason about noncriticality of distance, as follows:

Suppose $\mathrm{dist}_M(p,x)$ is large enough to satisfy the lemma, for some fixed small δ. Choose γ fixed and then γ_1, in the notation of the lemma, and choose t very large. Then $\mathrm{dist}_M(x,\gamma_1(t))$ is estimated from above, compared to the euclidean triangle with sides $\mathrm{dist}_M(p,x)$ and $\mathrm{dist}_M(\gamma_1(0),\gamma_1(t))=t$ and included angle δ. Choose a minimal geodesic γ_2 from x to $\gamma_1(t)$. Now consider the triangle formed by γ_2, $\gamma_1 \,|\, [0,t]$ and *any* minimal geodesic from p to x. Toponogov's theorem shows again that the angle at x must be larger than $\pi/2$; otherwise the third side could not have length as large as its actual length t. Thus the vector γ_2' at x satisfies the geodesic condition for noncriticality of $\mathrm{dist}_M(\,\cdot\,,p)$ at x. These arguments are given in [Gromov 1981a] and generalized to asymptotic nonnegativity of curvature in [Kasue 1988].

The relatively elementary argument just given already shows, when combined with Lemma 3.1 (noncritical Morse theory), the following striking fact:

THEOREM 3.3 (TOPOLOGICAL FINITENESS). *A complete noncompact manifold of nonnegative sectional curvature is diffeomorphic to the interior of a compact manifold-with-boundary.*

Historically, this result was discovered in the more involved context of the convex function considerations [Cheeger and Gromoll 1972]. These considerations are the subject of the next section.

4. Of the Distance from Infinity and Convexity

A more profound analysis of the structure of complete, noncompact, nonnegative curvature manifolds can be obtained by exploiting systematically a construction originally introduced by H. Busemann. This construction attaches to each ray $\gamma : [0,+\infty) \to M$ (parametrized by arclength) a function that we shall call B_γ, the Busemann function of γ. Intuitively, B_γ is a measure of how far a point is out toward infinity in the direction of γ. The precise definition is

$$B_\gamma(x) = \lim_{t\to+\infty} (t - \mathrm{dist}_M(x,\gamma(t))).$$

The function $t \mapsto (t - \mathrm{dist}_M(x,\gamma(t)))$ is monotone nondecreasing and bounded above by $\mathrm{dist}_M(x,\gamma(0))$ so that the limit exists and is finite-valued. Also, since each of the functions $t \mapsto t - \mathrm{dist}_M(x,\gamma(t))$ is globally Lipschitz continuous with Lipschitz constant 1, the function B_γ is also Lipschitz continuous with Lipschitz constant 1.

These considerations are independent of any curvature hypotheses. But in the presence of everywhere nonnegative curvature, the Busemann functions acquire a new virtue: they are necessarily convex. From the perspective of support functions and Toponogov's Comparison Theorem developed earlier, this convexity is almost immediate, as will now be shown.

Nonnegative curvature implies, either by Toponogov's Comparison Theorem or by a direct second variation argument, that the function $x \mapsto -\operatorname{dist}_M(x, \gamma(t))$ has second derivatives, along arclength geodesics in a neighborhood of a fixed x_0, whose value is at least $-2/\operatorname{dist}_M(x_0, \gamma(t))$. This is true in the sense of support functions even if the function $\operatorname{dist}(\,\cdot\,, \gamma(t))$ does not have two derivatives literally; that is, along such a geodesic $\theta(s)$, the function $\delta s^2 - \operatorname{dist}_M(\theta(s), \gamma(t))$ is convex for $\delta \geq \operatorname{dist}_M(x_0, \gamma(t))$. It follows that the same support function second derivative estimate is true of $t - \operatorname{dist}_M(x, \gamma(t))$. By taking the limit, one immediately derives the convexity of B_γ.

An elementary argument shows that, if $f : M \to \mathbb{R}$ is a (geodesically) convex function, then f is necessarily locally Lipschitz continuous. (This and subsequent remarks on general convex function theory are independent of curvature hypotheses.) Thus the whole machinery of noncriticality can be applied. And in this regard, convex functions have an extraordinary property:

Theorem 4.1 (Noncriticality of convex functions [Greene and Shiohama 1981a; 1981b]). *If $f : M \to \mathbb{R}$ is a convex function on a complete Riemannian manifold and if x is a point of M such that $f(x) > \inf_M f$, then x is a noncritical point for f.*

Outline of proof. By connecting x via a geodesic to a point y with $f(y) < f(x)$, one sees that $\operatorname{dist}_M(x, f^{-1}((-\infty,\, f(x) - \delta]))$ goes to 0 as $\delta \to 0^+$. The set $f^{-1}((-\infty,\, f(x) - \delta])$ is convex. From this convexity, one sees that there is a unique shortest connection from each point near x to this set when $\delta > 0$ is small enough. Uniqueness shows that this shortest connection varies continuously with the variation of x, near x. The tangent vectors to these geodesics satisfy the vector field condition for there to be a subgradient for f at x. (See [Greene and Shiohama 1981a; Greene and Shiohama 1981b] for this argument in detail.) ☐

It is actually most advantageous to apply these noncriticality considerations not to the individual Busemann functions but to the function $B(x) = \sup_\gamma B_\gamma(x)$, where the supremum is taken over all rays γ with fixed $\gamma(0) = p$. The supremum of convex functions is convex, so when M has everywhere nonnegative sectional curvature, the function B is again convex. Moreover, the function B has compact sublevel sets, that is, $B^{-1}((-\infty, \alpha])$ is compact for each $\alpha \in \mathbb{R}$.

The proof is again a variant of the basic ray construction: If $\{x_j\}$ is a sequence converging to infinity, that is, $\lim \operatorname{dist}(p, x_j) = +\infty$, if $B(x_j) \leq \alpha$ for all j and if γ_j is a sequence of minimal, arclength-parameter geodesics, from p to x_j for each j, then B is bounded above on γ_j, by $\max(B(p), B(x_j)) \leq \max(B(p), \alpha)$, by the convexity of B. Let γ be a ray which is a limit of a subsequence of the γ_j

(such a γ exists). Along γ, B is bounded by $\max(B(p), \alpha)$, by continuity. But B_γ is unbounded on γ, hence so is B. This contradiction completes the proof.

These observations give another proof for Theorem 3.3 on topological finiteness: M is diffeomorphic to the interior of a compact manifold-with-boundary homeomorphic to $B^{-1}((-\infty, \alpha])$, for any $\alpha > \inf_M B$. This follows by combining Lemma 3.1 (noncritical Morse theory), Theorem 4.1, and the convexity and properness of B.

But the function B enables one to refine this picture further. First, it is not hard to see that, if $\beta < \alpha$, then

$$B^{-1}((-\infty, \beta]) = \{x : B(x) \leq \alpha,\, \operatorname{dist}_M(x, B^{-1}([\alpha, +\infty))) = \alpha - \beta\}.$$

It follows that if $\beta_0 = \inf_M B = \min_M B$, then $B^{-1}(\{\beta_0\})$ is a compact convex subset of M with empty interior. This set thus lies in a totally geodesic submanifold of M, and indeed is the closure of (an open subset U of) such a submanifold. The proof of this is almost a copy of the corresponding structural result for compact, convex subsets of $\mathbb{R}^n$. The difference here is that the closure of the submanifold may equal the submanifold, that is, the submanifold may have no boundary. This can happen in $\mathbb{R}^n$ only if the submanifold is a single point.

Given such a compact convex subset C in M, equal to the closure of a totally geodesic submanifold U, one can consider the *δ-push-in* $C_\delta = \{x \in U : \operatorname{dist}_M(x, C - U) \geq \delta\}$, provided that $C - U \neq \varnothing$.

It is not hard to see, using support functions as usual, that the function $x \mapsto -\operatorname{dist}(x, C - U)$ is convex on U. In particular, the push-ins C_δ are convex, for each $\delta > 0$. There is a maximal δ for which C_δ is nonempty. And the same structural result for this C_δ, which is necessarily of lower dimension, enables one to do the push-in construction again in a lower-dimensional totally geodesic submanifold.

Since dimension drops at each stage, this process must eventually terminate with a maximal push-in that is a totally geodesic submanifold without boundary. Clearly, M has the homotopy type of this submanifold. But in fact, standard topological neighborhood constructions [Rushing 1973], together with the noncriticality of convex functions already discussed, show that M is diffeomorphic to a tubular neighborhood of the submanifold δ. Thus one obtains what is usually called the Soul Theorem:

THEOREM 4.2 (STRUCTURE OF NONNEGATIVELY CURVED MANIFOLDS [Cheeger and Gromoll 1972; Poor 1974]). *If M is a complete, noncompact manifold of everywhere nonnegative curvature, then there is a compact totally geodesic submanifold (without boundary) S of M such that M is diffeomorphic to a tubular neighborhood of S. In particular, M is diffeomorphic to the total space of a vector bundle over a compact manifold of nonnegative curvature.*

The submanifold S is usually called the *soul* of M, an unattractive but apparently permanent piece of terminology. The theorem was largely established in [Cheeger and Gromoll 1972], but the diffeomorphism statement was not obtained in full generality until [Poor 1974]. Of course, here and throughout the discussion, the fact has been used repeatedly that a totally geodesic submanifold of a manifold of nonnegative sectional curvature has itself nonnegative sectional curvature.

In the particular situation of complete noncompact manifolds of nonnegative curvature, it is not necessary to appeal to the topological generalities referred to before the statement of theorem. By using the noncritical Morse theory for nonsmooth functions that is a main theme in the present article, one can deduce more directly that such a manifold M is diffeomorphic to a tubular neighborhood of a soul S of M. For this, one needs only that the function $x \mapsto \operatorname{dist}(x, S)$ is noncritical at every point $x \notin S$. For then, by the smooth approximation of subgradient flow as developed in [Greene and Shiohama 1981a; 1981b], for example, one obtains a diffeomorphism of M onto $\{x : \operatorname{dist}(x, S) < \varepsilon\}$, for any $\varepsilon > 0$. For ε small enough, this latter set is a tubular neighborhood, and the proof is complete.

To see that $x \mapsto \operatorname{dist}(x, S)$ is noncritical at each $x \notin S$, fix such an x and consider a minimal arclength parameter geodesic γ from x to S, so that $\operatorname{length}(\gamma) = \operatorname{dist}(x, S)$, while γ starts at x and terminates at a point in S. Suppose first that $B(x) > \min_M B$, where B is the supremum of the Busemann functions, as before. The convexity of B implies that the rate of decrease of B along γ is at least $(B(x) - \min B)/\operatorname{length}\gamma$ at the point x. Moreover, the convexity of B also implies that the set of unit vectors in T_xM along which B decreases at least at that rate is contained in a closed convex cone in T_xM that lies in an open half-space. In particular, there is a unit vector $u \in T_xM$ and an $\alpha > 0$ such that the angle from u to any such geodesic γ is $\geq \pi/2 + \alpha$. Hence, by the logic already discussed in comparing the geometric noncriticality of [Grove and Shiohama 1977] with the analytic noncriticality idea used here, it follows that $\operatorname{dist}(\,\cdot\,, S)$ is noncritical (in the analytic sense) at x.

If $B(x) = \min_M B$ but still $x \notin S$, then one can apply the same reasoning to deduce the noncriticality of $\operatorname{dist}(\,\cdot\,, S)$ at x. In this case, one works not with B but with the relevant convex function associated to the push-in that takes x to the next-lower-dimensional totally convex set that is the next stage in the construction of the soul. By total convexity, all minimal connections γ from x to S lie in a totally convex set containing x and S on which the convex function for the push-in is defined. So the argument goes exactly as before.

Striking as this structural result is, it leaves a number of substantial questions unanswered. One of them is, which vector bundles over compact manifolds S of nonnegative sectional curvature actually do admit complete metrics of nonnegative sectional curvature on their total spaces? Trivial bundles obviously do: one just uses the product metric on $S \times \mathbb{R}^n$. In practice, it is not easy to find cases where one can see that a certain vector bundle does not occur. This makes the following results, proved in [Özaydin and Walschap 1994], particularly intriguing: The total space of a rank k vector bundle over a compact flat manifold M admits a complete metric of nonnegative sectional curvature if and only if E admits a complete flat metric, and also if and only if E is diffeomorphic to $\mathbb{R}^n \times_{\pi_1(M)} \mathbb{R}^k$, where $\pi_1(M)$ acts on $\mathbb{R}^n$ by covering transformations (of M) and on $\mathbb{R}^k$ by an orthogonal representation. This latter condition is in turn equivalent to E being diffeomorphic to the total space of a vector bundle of rank k over M that admits a flat Riemannian connection. From this, one deduces that the total space of a rank-2 oriented vector bundle over a compact flat manifold admits a complete metric of nonnegative sectional curvature if and only if its rational Euler class vanishes. This provides many examples of vector bundles over flat manifolds the total space of which does not admit a complete metric of nonnegative sectional curvature—for example, among oriented rank-2 bundles over the n-torus, only the trivial bundle has a complete metric of nonnegative sectional curvature.

So far in this section, attention has been concentrated on B, the supremum of the Busemann functions over all rays from a given point, for the logical reason that in general relatively little information can be derived by considering a single Busemann function. But there is one situation wherein much information can be derived about a single Busemann function: Suppose that a complete Riemannian manifold M contains a line $\gamma : (-\infty, +\infty) \to M$, that is a curve with

$$\operatorname{dist}(\gamma(t_1),\ \gamma(t_2)) = |t_1 - t_2|$$

for all $t_1, t_2 \in \mathbb{R} = (-\infty, +\infty)$. Associated to this situation are two Busemann functions: B_+, the Busemann function of the ray $\gamma|[0, +\infty)$, and B_-, that of the ray $t \mapsto \gamma(-t)$, for $t \in [0, +\infty)$.

It follows from the triangle inequality that $B_+ + B_- \leq 0$ everywhere on M and that $B_+ + B_- = 0$ at points of γ. This is not conspicuously instructive in complete generality. But if M is supposed to have nonnegative Ricci curvature, second variation arguments already discussed show that Busemann functions on M are subharmonic in the support sense (and even convex, under the more restrictive assumption of nonnegative sectional curvature).

Now the maximum principle holds for support-subharmonic functions, as is easy to see. Thus the fact that $B_+ + B_-$ is nonnegative on M and vanishes on γ implies that it vanishes in fact on all of M, if B_+ and B_- are subharmonic—e.g., when M has everywhere nonnegative Ricci curvature. Then B_+ is support-harmonic (since it is subharmonic and its negative B_- is also subharmonic). So, by the smoothness principle of Section 2, B_+ is a C^∞ harmonic function. (When M has nonpositive sectional curvature, the weaker and more elementary smoothness principle for functions linear along geodesics suffices here.)

Since B_+ is then C^∞, it has a gradient everywhere, of length 1. The gradient has length no more than 1 because B_+ is Lipschitz continuous with constant 1.

To see that the gradient has length at least 1 at a given $x \in M$, choose a sequence of minimal, arclength parameter geodesics γ_j from x to points $\gamma(t_j)$ with $t_j \to +\infty$ and unit vectors γ_j at x converging to a (unit) vector v. Then B_+ can easily be seen to increase at unit rate along $t \mapsto \exp_x(tv)$, $t \geq 0$.

Moreover, B_+ being harmonic implies that its level surfaces, which are smooth hypersurfaces, have mean curvature 0. Since they are "equidistant", it follows easily that they are totally geodesic hypersurfaces. One then sees immediately that M must be isometric to

$$\{x \in M : B_+(x) = 0\} \times \mathbb{R},$$

the isometry being the map that associates to each (x, t) the flow to time t along the integral curve of grad B_+ emanating from x at $t = 0$.

This gives the Splitting Theorem of [Cheeger and Gromoll 1971] for manifolds of nonnegative Ricci curvature (and, by the easier form of the arguments, Toponogov's Splitting Theorem [Toponogov 1964] for manifolds of nonnegative sectional curvature): If a complete n-dimensional manifold of nonnegative Ricci curvature contains a line, it is isometric to the product of an $(n-1)$-manifold and the real line $\mathbb{R}$.

5. Of Distance-Nonincreasing Retractions and Manifolds with Positive Curvature at One Point

If C is a closed subset of a complete Riemannian manifold M, then, for each $p \in M$, there is a point p' in C that is as close to p as possible, that is,

$$\operatorname{dist}(p, p') = \operatorname{dist}(p, C) := \inf_{q \in C} \operatorname{dist}(p, q).$$

If C is a closed, convex set in $\mathbb{R}^n$, then, for each $p \in \mathbb{R}^n$, there is exactly one such closest point p' in C. In this case, one can define a retraction R of $\mathbb{R}^n$ onto C by setting $R(p) = p'$. This retraction R has the property of being distance-nonincreasing in the sense that, for all $p_1, p_2 \in \mathbb{R}^n$,

$$\operatorname{dist}(R(p_1), R(p_2)) \leq \operatorname{dist}(p_1, p_2).$$

For complete Riemannian manifolds, the situation is more complicated. The "closest point" p' may not be unique: for example, the north pole is equidistant from every point of the boundary of the southern hemisphere (which is convex). However, if C is a closed, convex set in a complete Riemannian manifold, there is a neighborhood U of C such that, for each $p \in U$, there is a unique closest point p' in C; a proof is given in [Greene and Shiohama 1981a], where the problem was investigated in detail.

Even if two points p_1, p_2 have unique closest points p'_1 and p'_2, and if C is convex, it may be that $\operatorname{dist}(p'_1, p'_2) > \operatorname{dist}(p_1, p_2)$. For instance, this happens when C is the southern hemisphere and p_1, p_2 are points of the same latitude in the northern hemisphere (excluding the equator and pole).

This latter example, if looked at in some detail, however, strongly suggests that there is some kind of control over closest-point retractions onto convex sets. If p_1, p_2 are on the parallel of latitude ε north, and are separated in longitude by θ, the law of spherical cosines gives

$$\cos \operatorname{dist}(p_1, p_2) = \sin^2 \varepsilon + cos^2 \varepsilon \cos \theta,$$

where we have chosen units making the radius of the sphere equal to 1. On the other hand,

$$\cos \operatorname{dist}(p_1', p_2') = \cos \theta.$$

Since $\cos \varepsilon = 1 - \frac{1}{2}\varepsilon^2 + \dots$ and $\sin \varepsilon = \varepsilon - \frac{1}{6}\varepsilon^3 + \dots$, one sees that the retraction $R : p_i \mapsto p_i'$, as before, is *almost* distance-nonincreasing, where almost means up to an error that is second order in the distance ε of p_1, p_2 from C.

The historic concept of "infinitesimals of higher order", or Duhamel's Principle as it is often called, now suggests that, if some similar second-order statement holds for Riemannian manifolds in general, then it should be possible to construct distance-nonincreasing retractions onto (sub)level sets of convex functions. In outline, with p_1, p_2 in the same level, say α, one projects p_1, p_2 successively onto closely-spaced lower levels via closest-point mappings, until one reaches some fixed lower level β. These successive projections make at most second-order increases in distance. So by taking the limit with the number of intermediate levels going to infinity, one obtains a retraction that is distance-nonincreasing. The "errors" have vanished in the limit, having been infinitesimals of higher order.

This old-fashioned, almost eighteenth-century, way of thinking can in fact be made easily into a precise proof of the following result, first obtained by Sharafutdinov [1977] for retraction on the soul of a manifold of nonnegative curvature (compare [Greene and Shiohama 1981a]):

Distance-nonincreasing retraction theorem. *If M is a complete Riemannian manifold, and if $f : M \to \mathbb{R}$ is a convex function, then, for each $\alpha \in f(M)$, there is a retraction*

$$R : M \to f^{-1}((-\infty, \alpha])$$

satisfying

$$\operatorname{dist}_M(R(x_1),\, R(x_2)) \le \operatorname{dist}_M(x_1, x_2)$$

for all $x_1, x_2 \in M$.

To make the proof of this result more explicit, note first that it is enough to construct, for a given fixed $\beta > \alpha$, a distance-nonincreasing retraction $R_{\beta,\alpha}$ from $f^{-1}([-\infty, \beta])$ onto $f^{-1}([-\infty, \alpha])$. Then the required R on M can be obtained as a composition of $R_{\alpha+1,\alpha}$, $R_{\alpha+2,\alpha+1}$, etc. To construct $R_{\beta\alpha}$ (when f is proper), note that for all sufficiently large positive integers n, there are unique closest points in $f^{-1}([-\infty,\, \alpha + (\beta-\alpha)i/n])$ to points in $f^{-1}([-\infty,\, \alpha + (\beta - \alpha)(i +$

$1)/n]$), for $i = 0, 1, \ldots, n-1$. This follows from the existence of a unique closest point for points in a small neighborhood of a convex set, as noted earlier. This construct defines a retraction R of $f^{-1}([-\infty, \beta])$ onto $f^{-1}([-\infty, \alpha])$ by composition of the successive closest-point maps. Each of these closest-point maps is distance-nonincreasing modulo an error of order n^{-2}; this follows from a geometric argument that will be made explicit momentarily. A (sub)sequence of the R_n converges uniformly to a retraction that is distance-nonincreasing: the error terms are n in number and $O(n^{-2})$, so that their contribution vanishes in the limit.

Two further ingredients are needed to complete the geometric part of this argument, that is, to establish that projections onto the lower sublevels are distance-nonincreasing up to errors of order n^2. The first ingredient is to note an angle property of projection: Suppose that C is a compact (totally) convex set, and that x_1 and x_2 are points not in C but close enough to C to have unique closest points in C, say y_1 and y_2. Let γ_1 be minimal from x_1 to y_1, γ_2 from x_2 to y_2, and γ_3 from y_1 to y_2. Then the angle at y_1 between γ_1 and γ_3 (that is, $\angle x_1 y_1 y_2$) and the corresponding angle at y_2 are both $\geq \pi/2$. This is an obvious consequence of the first variation formula and the fact that $\gamma_3 \subset C$: If, say, $\angle x_1 y_1 y_2$ were acute, then moving along γ_3 from y_1 towards y_2 would give points (in C) closer to x_1 than is y_1, whereas y_1 is the closest point in C to x_1. In an obvious sense, this angle estimate shows that $\operatorname{dist}(x_1, x_2) \geq \operatorname{dist}(y_1, y_2)$ up to terms of second order in $\operatorname{dist}(x_1, C)$ and $\operatorname{dist}(y_1, C)$. (Here $\operatorname{dist}(x_2, y_2)$ is bounded uniformly if C is compact; otherwise, in case one wants to consider C only closed, convex, not necessarily compact, one still has a bound on $\operatorname{dist}(y_1, y_2)$ in terms of $\operatorname{dist}(x_1, x_2)$ and $\operatorname{dist}(x_1, C)$ and $\operatorname{dist}(x_2, C)$, which suffices.)

The second ingredient needed is the observation that, supposing the sublevel $f^{-1}((-\infty, b])$ to be compact and a to be greater than $\min_M f$, there is a constant C_0 independent of n such that

$$f(x) \in [a + i(b-a)/n,\ a + (b-a)(i+1)/n] \quad \text{for } i = 0, 1, \ldots, n-1 \implies \operatorname{dist}(x, f^{-1}((-\infty,\ a + (b-a)i/n])) \leq C_0/n.$$

This estimate shows that "second-order" in terms of distance to the closest-point projection, as in the previous paragraph, implies second-order in the sense of order $1/n^2$. To establish this estimate, choose a number α such that $\alpha < a$ but $\alpha > \min_M f$. For x such that $f(x) \in [a + (b-a)i/n,\ a + (b-a)(i+1)/n]$, choose $x_0 \in f^{-1}((-\infty, \alpha])$ such that

$$\operatorname{dist}(x, x_0) = \operatorname{dist}(x, f^{-1}((-\infty, \alpha])).$$

And choose a minimal, arclength parameter geodesic $\gamma : [0, D_x] \to M$ with $\gamma(0) = x_0$ and $\gamma(D_x) = x$.

Clearly $f(\gamma(t)) > \alpha$ for $t > 0$. Let t_0 be the smallest positive number such that $f(\gamma(t_0)) = a$. Now the function $t \mapsto f(\gamma(t))$ is convex. It follows that f is strictly increasing on $[0, D_x]$, and that if t_1 is the unique number in $[0, D_x]$ such that $f(\gamma(t_1)) = a + (b-a)i/n$, then

$$\frac{D_x - f(\gamma(t_1))}{D_x - t_1} \geq \frac{f(\gamma(t_0)) - f(\gamma(0))}{t_0 - 0}.$$

Thus

$$\begin{aligned}\operatorname{dist}(x, f^{-1}(-\infty, a+(b-a)i/n]) \leq D_x - t_1 &\leq \frac{t_0}{a-\alpha}\big(f(x) - (a+(b-a)i/n)\big)\\ &\leq \left(\frac{t_0}{a-\alpha}\right)\frac{1}{n} \leq \frac{d_\alpha}{a-\alpha}\,\frac{1}{n},\end{aligned}$$

where $d_\alpha = \inf\{\operatorname{dist}(p,q) : f(p) = a,\ f(q) = \alpha\}$. Thus the required estimate holds, with $C_0 = d_\alpha(a-\alpha)$.

Retraction that is distance-nonincreasing onto $a = \min f$ follows by a limiting argument from the case of $a > \min f$. Also, if the sublevels are not compact, similar estimates hold on compact subsets and the desired constructions can be carried out globally by patching arguments, which are omitted in the interests of brevity.

A variant of this argument can be used to show that if $B > \alpha > \inf_M f$ then $f^{-1}([-\infty, \beta])$ is homeomorphic to $f^{-1}([-\infty, \alpha])$, and also that $f^{-1}((-\infty, \alpha))$ is homeomorphic to M. One does not map a point p in $f^{-1}([\alpha + (\beta-\alpha)i/n,$ $\alpha + (\beta-\alpha)(i+1)/n])$ to its closest point in $f^{-1}([-\infty,\ \alpha+(\beta-\alpha)i/n])$. Rather, one takes the point p to the intersection with $\{x : f(p) = \alpha + (\beta-\alpha)i/n\}$ of the geodesic from p to the closest point to p in $f^{-1}(([-\infty,\ \alpha+(\beta-\alpha)(i-1)/n]))$—or, if $i = 0$, to $f^{-1}(([-\infty,\ \alpha-\varepsilon]))$, for some small $\varepsilon > 0$. This construction gives a sequence of "push-downs" that can be patched together to exhibit a product structure on $f^{-1}([\alpha, \beta])$, making it homeomorphic to $\{x \in M : f(x) = \alpha\} \times [\alpha, \beta]$. (See [Greene and Shiohama 1981a] for details, and [Greene and Shiohama 1981b] for a proof of the stronger result where "diffeomorphism" is substituted for "homeomorphism".) This process serves as a geometric substitute for the more analytic viewpoint of noncriticality of convex functions away from their minimum sets, discussed in Section 3.

The soul S of a complete manifold M of nonnegative sectional curvature is obtained as successive push-downs to the minimum set of convex functions, as explained in Section 4. Thus one obtains the following corollary (established in [Sharafutdinov 1977]) of the result on distance nonincreasing retractions:

Retraction onto the soul. *If M is a complete Riemannian manifold with everywhere nonnegative sectional curvature, and if $S \subset M$ is a soul of M, then there is a distance-nonincreasing retraction $R : M \to S$.*

The retraction R is, as noted, obtained as a composition of retractions that are themselves limits of compositions of other retractions. This might seem to be a rather uncontrolled, even uncontrollable, construction, from the geometric point of view. Thus, the following result of Perelman [1994] is more than a little startling:

Rigidity of retraction onto the soul. *If $R : M \to S$ is a distance-nonincreasing retraction of a complete manifold M of nonnegative curvature onto a soul S, then $R(\exp_x v) = x$ for each $x \in S$ and $v \in T_xM$ with $v \perp T_xS$.*

The exponential map on the normal bundle of S need not be a diffeomorphism onto M, even though M is indeed diffeomorphic to this normal bundle (see Section 3). For instance, if S is a point, so that M is diffeomorphic to $\mathbb{R}^n$, it may not be the case that M has any pole (recall that $x \in M$ is a *pole* if $\exp_x$ is a diffeomorphism). Even so, Perelman's rigidity result can well be thought of as saying that R is a one-sided inverse of exp of the normal bundle of S.

If M is a complete Riemannian manifold of everywhere positive sectional curvature, the supremum B of the Busemann functions is actually strongly convex in the support-function sense. That is, for each x in M and for σ a C^∞ function in a neighborhood of x with positive definite covariant differential at x, there is an $\varepsilon > 0$ such that $B - \varepsilon\sigma$ is convex in a neighborhood of x. The minimum set of a strongly convex function (in the support-function sense, which applies to not necessarily smooth functions) is either empty or consists of a single point. Thus the soul S of M is a point and M is diffeomorphic to $\mathbb{R}^n$, where $n = \dim M$. Homeomorphism to $\mathbb{R}^n$ was proved by Gromoll and Meyer [1969], prior to the general structure theorem for nonnegative curvature, and indeed this work introduced already the basic results about the suprema of the Busemann functions and so on. The explicit proof of diffeomorphism to $\mathbb{R}^n$, as opposed to just homeomorphism, came in [Poor 1974] (this is of course relevant only in dimensions 3 and 4). The same result was proved in [Greene and Wu 1976] by a particularly direct process: we first proved that the function B can be smoothed, that is, approximated by a strongly convex C^∞ function, which is also proper. That M is diffeomorphic to $\mathbb{R}^n$ then follows from standard smooth Morse theory techniques.

It was natural to ask whether the conclusions that S is a point and hence that M is diffeomorphic to $\mathbb{R}^n$ hold under the weaker hypotheses that M has nonnegative curvature everywhere and all sectional curvatures positive at a single point. This was known to be true in two particular cases: in dimension 2 [Cohn-Vossen 1935], and for manifolds M that arise as the boundaries of convex bodies in $\mathbb{R}^{n+1}$ (see the classification in [Busemann 1958]). Of course, for $n > 2$, it is atypical for a general Riemannian manifold to be isometric to a hypersurface, even locally, since local isometric embedding requires, generically, $\frac{1}{2}n(n+1)$ dimensions at least (compare [Gromov and Rohlin 1970]). In any case, it was conjectured by Cheeger and Gromoll [1972] that sectional curvatures nonnegative everywhere and all positive at one point (on a complete, noncompact manifold) implies diffeomorphism to euclidean space.

Only very limited special cases of this conjecture were confirmed in the next twenty years: for example, when S has codimension 1 [Cheeger and Gromoll 1972] or 2 [Walschap 1988], and some cases involving detailed curvature assumptions [Elerath 1979]. Thus Perelman's proof of the conjecture, presented for the first time at the MSRI Workshop of which this volume is a record, caused excitement. In fact the solution follows from this general result [1994]:

FLAT RECTANGLES IN NONNEGATIVE CURVATURE. *Let M be a complete noncompact manifold of nonnegative curvature with soul S of positive dimension. Let γ be a geodesic in S, and let N be a parallel vector field along γ normal to S. Then the curves γ_t defined by*

$$\gamma_t(u) = \exp_{\gamma(u)} tN(u),$$

for $t \geq 0$, are geodesics in M and form a flat totally geodesic two-dimensional submanifold.

COROLLARY 5.1 [Perelman 1994]. *A complete, noncompact, n-dimensional manifold with everywhere nonnegative sectional curvature and with all sectional curvatures positive at some point is diffeomorphic to $\mathbb{R}^n$.*

The corollary follows from the theorem because every point p of M arises as $\exp_x tN$, where $x \in S$ and $N \perp T_xS$. This is easy to see, e.g., by taking for x a closest point in S to p and for N the tangent vector at x to a minimal-length geodesic from x to p. If S has positive dimension the theorem provides a zero sectional curvature at p. So if all sectional curvatures at some p are positive, S must be zero-dimensional, that is a point (S is convex, and therefore connected). Hence M is diffeomorphic in that case to $\mathbb{R}^n$.

The arguments used by Perelman to prove the rigidity theorem and the flat two-plane theorem are ingenious, but purely geometric and, in a sense, surprisingly elementary. The starting point is to define a function $f(r)$ as follows: if $R : M \to S$ is a distance-nonincreasing retraction, set

$$f(r) = \max \operatorname{dist}(x, R(\exp_x rv)),$$

where $x \in S$ and $v \in T_xM$ is orthogonal to S. Then, by a geometric argument, one shows that the upper (one-sided) derivative of f is nonnegative. Since f is clearly Lipschitz continuous, f is the integral of its derivative up to an additive constant. Since $f(0) = 0$ while $f \geq 0$ by definition, $f(r) \equiv 0$ for $r \geq 0$. This establishes the rigidity. The flat two-plane result then follows using comparison results applied to "rectangles" obtained by exponentiating the parallel vector field N along γ, applying R, and using the rigidity result.

It was also shown in [Perelman 1994] as part of the proofs of the results already quoted that Sharafutdinov's retraction R was a Riemannian submersion of class C^1. Recently, this submersion has been shown to be of class C^∞ by Guijarro [a].

6. Of How Little Curvature a Manifold of Nonnegative Curvature Can Have

In Section 1, the question was explored of how much positive curvature a noncompact manifold of nonnegative curvature could have. That there was some restriction was obvious, since if there were too much positive curvature, the manifold would close up, ceasing to be noncompact. But at first sight, there seems to be no restriction in the other direction, no lower bound on how much positivity of curvature can occur. After all, the most familiar examples of open manifolds of nonnegative curvature, euclidean spaces, have no positive curvatures at all; they are flat. It is thus rather surprising that, if one supposes in advance that the metric is not flat, there is a precise sense, in many cases, in which the manifold cannot be arbitrarily close to flat.

The first results of this type were discovered only relatively recently; they are not extensions of results from the classical period of Riemannian geometry, in the late nineteenth and early twentieth centuries. The absence of any such results classically is probably just a consequence of the fact that for surfaces there are no results of the type. It is easy, for instance, to construct complete metrics (from surfaces of revolution) on $\mathbb{R}^2$ that have zero curvature outside some compact set, nonnegative curvature everywhere, and positive curvature somewhere. The corresponding situation for metrics on $\mathbb{R}^n$, for $n \geq 3$, cannot occur, as we shall see momentarily. And this nonoccurrence is the basic instance of the restrictions on near-flatness without total flatness with which we shall be concerned.

The first results of this general type were actually obtained for complete, simply connected Kähler manifolds of nonpositive curvature and "faster-than-quadratic curvature decay", in the sense that, for some $C > 0$ and $\varepsilon > 0$,

$$\sup |K(\sigma)| \leq C r^{-(2+\varepsilon)},$$

where the supremum of absolute value of sectional curvature is taken over all two-planes σ at distance r from a fixed point. Such manifolds had been proposed as objects of study in [Greene and Wu 1977], where, in particular, it was conjectured that they should be necessarily biholomorphic to $\mathbb{C}^n$. This conjecture was proved in [Siu and Yau 1977]. But except in complex dimension 1, no examples were known except the standard metric on $\mathbb{C}^n$. And in [Mok et al. 1981], it was shown that indeed such manifolds had to be isometric to $\mathbb{C}^n$ if their complex dimension n were ≥ 2. (The biholomorphism theorem was extended in the same paper to allow small amounts of positive curvature, so that then nonflat examples occur). This work on Kähler geometry motivated H. Wu and myself to consider the corresponding, more general Riemannian situation to be discussed now.

The most basic and prototypical Riemannian-geometric result has to do with manifolds that are flat outside some compact set:

THEOREM 6.1 [Greene and Wu 1982]. *Let M be a complete noncompact manifold such that* (a) *there is a compact set K such that all sectional curvatures are* 0 *on $M - K$, and* (b) *M is connected at infinity and simply connected at infinity. Then there are compact sets K_1 in M and K_2 in $\mathbb{R}^n$ with $M - K_1$ isometric to $M - K_2$.*

COROLLARY 6.2 [Greene and Wu 1982; Greene and Wu 1993]. *If a manifold satisfying the hypotheses of the Theorem also has the property of everywhere nonnegative Ricci curvature, it is isometric to $\mathbb{R}^n$.*

COROLLARY 6.3 [Greene and Wu 1982]. *If a complete, simply connected Riemannian manifold of dimension $n \geq 3$ has everywhere nonpositive sectional curvature and has zero curvature outside some compact set, it is isometric to $\mathbb{R}^n$.*

The first corollary follows from the theorem by considering a (large) region in M containing K_1, the boundary of which is isometric to the boundary of a cube in $\mathbb{R}^n$. Identifying opposite faces gives a manifold of dimension n with nonnegative Ricci curvature and with first homology of rank n. Such a manifold must be flat, by the Bochner technique, so M itself is flat.

The second corollary follows from the Theorem by noting that M has exactly euclidean volume growth, that is,

$$\lim_{r \to +\infty} \operatorname{vol} B(p,r) / \operatorname{vol}_e B(r) = 1$$

for each $p \in M$, where $\operatorname{vol}_e B(r)$ is the volume of the euclidean ball of radius T. Then M must be flat by the Bishop volume comparison result [Bishop and Crittenden 1964]: if any sectional curvature of M at a point p, say, were negative, the volume growth limit would be greater than 1 from p.

It is an important fact that no analogue of the first corollary holds for nonpositive Ricci curvature: There are compactly supported perturbations of the standard euclidean metric on $\mathbb{R}^n$, for $n \geq 3$, which have Ricci curvature nonpositive everywhere and negative somewhere [Lohkamp 1992].

The proof of the Theorem on manifolds flat outside a compact set is obtained by considering the locally isometric "developing map" of M outside a compact set to $\mathbb{R}^n$. Using the convexity results of [Greene and Wu 1973; 1974], one can take the compact set to have C^∞ boundary and to be convex in M. Simple connectivity gives a well-defined developing map which maps the boundary of the compact convex set in M to an immersed, locally convex hypersurface in $\mathbb{R}^n$. Since $n \geq 3$, this immersion is in fact an embedding [Sacksteder 1960; van Heijenoort 1952], and the consideration of exterior normal maps gives an isometric map from the exterior of this compact convex hypersurface in $\mathbb{R}^n$ to the exterior of the corresponding set in M. This technique can be extended to classify the possible structures of manifolds flat outside a compact set even when the manifolds are not assumed simply connected at infinity ([Greene and Wu 1993]; this classification was in fact obtained earlier by a different method in [Schroeder and Ziller 1989]).

It is natural to try to replace the condition of flatness outside a compact set by some condition of rapid decay of curvature to zero with increasing distance. The Kähler-geometric situation suggests that faster than quadratic decay is the appropriate choice. So does the relationship between faster-than-quadratic curvature decay and quasi-isometry of the exponential map established in [Greene and Wu 1979]. In [Greene and Wu 1982], the faster-than-quadratic-decay condition was analyzed as a substitute for flat-outside-a-compact-set in the case of manifolds admitting a pole.

In summary form, it was proved in [Greene and Wu 1982] that a manifold with a pole, with faster than quadratic curvature decay, and with sectional curvature of one sign (either everywhere ≤ 0 or everywhere ≥ 0) was necessarily flat, if the dimension of the manifold was ≥ 3, except when the dimension is 4 or 8 and the curvature is nonnegative.

These results were called "gap theorems" because in effect they showed the existence of a gap between flat $\mathbb{R}^n$ and other metrics of signed curvature on $\mathbb{R}^n$: such metrics cannot be too close to the flat metric.

The basic proof technique for these results of [Greene and Wu 1982] is to estimate the $(n-1)$-dimensional volume of distance spheres with a view to showing that M has euclidean volume growth. Then the flatness follows from the Bishop volume comparison [Bishop and Crittenden 1964], as already mentioned. For curvature ≤ 0, one wishes to estimate volume from above. One can estimate the sectional curvature of distance spheres from below with a positive lower bound depending on the radius, using comparison methods such as the Hessian comparison of [Greene and Wu 1979]. The volume estimate above follows. (This argument, simpler than the ones used in [Greene and Wu 1982], was communicated to the author by M. Gromov.) But for curvature ≥ 0, one seeks estimates of $(n-1)$-dimensional volume of the distance spheres from below, while the information available directly from comparison theory is an upper (and lower) curvature bound on the spheres' intrinsic metrics. In general, such curvature information fails to yield a lower bound on volume, as the phenomenon of "collapse with bounded curvature" shows (e.g., Berger spheres).

In odd dimensions one does get a lower volume estimate on the (even-dimensional) distance spheres by means of the generalized Gauss–Bonnet formula of [Allendoerfer and Weil 1943]. A bound on the absolute value of sectional curvature provides a bound on the generalized Gauss–Bonnet integrand and, since the integral must be the Euler characteristic of the even-dimensional sphere, namely 2, a lower volume bound follows. This method was introduced in [Greene and Wu 1982] for the odd-dimensional pole case. The same approach was used successfully in the case of manifolds not necessarily diffeomorphic to $\mathbb{R}^n$ (soul not a point) in [Eschenburg et al. 1989] to generate "gap theorems" in that situation again in odd dimensions, and to deal with the odd-dimensional case of diffeomorphism to $\mathbb{R}^n$, but with no pole.

When the manifold itself is even-dimensional, the estimation of distance sphere volume from below must use a different method. The idea used in [Greene and Wu 1982] is to replace the generalized Gauss–Bonnet integrand by a Gauss–Kronecker curvature computed relative to an almost parallel frame. (The almost parallel frame can be constructed in a neighborhood of a given (large) distance sphere because of the nearly-zero curvature of the manifold at large distances.) The Gauss–Kronecker integrand can be estimated above by comparison methods as before. Because the Gauss–Kronecker integrand is the determinant of the Gauss map relative to the almost parallel frame, its integral is, up to a normalization constant, equal to the degree of the Gauss map. Since the Gauss map is a classifying map for the tangent bundle of the sphere, the integral is bounded away from 0 (the degree being nonzero and in fact 1, *provided that the tangent bundle of* S^{n-1} *is nontrivial*, that is, provided that $n \neq 2$, 4, or 8). This method introduced in [Greene and Wu 1982] for the pole case was extended in [Drees 1994] to apply to curvature ≥ 0, no pole assumed. (The cases $n = 4$, 8 with a pole were treated in [Kasue and Sugahara 1987], but the method there does not generalize to the nonpole case.)

That $n = 4$ really is a special case was made a matter not just of restricted proof technique but also of concrete example in [Unnebrink a]. There a metric is constructed on $\mathbb{R}^4$ with cubic curvature decay and volume growth of cubic order, $\operatorname{vol} B(p, r) = O(r^3)$—an order of magnitude less than one might hope. Somewhat mysteriously, no example of a similar nature on $\mathbb{R}^8$ has yet been found.

Meanwhile, the author, P. Petersen, and S. Zhu have shown [Greene et al. 1994] that no such example can exist on any $\mathbb{R}^n$, for $n \neq 2, 4, 8$, with faster than quadratic curvature decay. And on $\mathbb{R}^n$, for $n = 4$ or 8, no such example exists with curvature decay faster than quartic; that is, if, for some C and $\varepsilon > 0$, one has

$$\sup |K(\sigma)| \leq C r^{-(4+\varepsilon)},$$

where the sup is over sectional curvatures at points of distance r from a fixed $p \in M$, then the volume growth is euclidean:

$$\lim_{r \to +\infty} \operatorname{vol} B(p, r) / \operatorname{vol}_{\mathrm{eu}} B(r) = 1.$$

(No hypothesis is made about signed curvature.) A corresponding result with faster-than-quartic decay also holds without topological hypotheses (of simple connectivity at infinity), the conclusion being in that case that the manifold admits a new metric, equivalent to the original one of fast curvature decay, with the new metric's sectional curvature being zero outside some compact set (the possibilities for metrics of that type having been already determined in [Schroeder and Ziller 1989] and [Greene and Wu 1993]).

7. Of Convergence and Behavior at Infinity

The rather specific proof techniques of the previous section can be put advantageously into a general picture via the concept of Gromov–Hausdorff convergence (hereafter, GH convergence). Suppose (M, g) is a complete, noncompact manifold with the property that

$$\liminf_{r \to +\infty} (r^2 \inf K(\sigma)) > -\infty,$$

where as usual the $\inf K(\sigma)$ means the infimum of sectional curvature $K(\sigma)$ over two-planes σ at distance r from a fixed base point p_0 in M. A manifold of faster-than-quadratic curvature decay has this property, for example. Then the family of Riemannian manifolds $\{(M, \lambda^{-1}g) : \lambda \geq 1, \lambda \in \mathbb{R}\}$ is precompact in the sense of GH convergence. (This precompactness property really requires only a corresponding estimate on Ricci curvature as a function of distance, but this greater generality would not be relevant here.) That is, every sequence $\{(M,\ \lambda_i^{-1}g) : \lim \lambda_i = +\infty\}$ has a subsequence that converges in the GH sense for pointed metric spaces, all the $(M,\ \lambda^{-1}g)$ being taken to have p_0 as base point. (See [Gromov 1981b] for the relevant convergence definitions.) The role here of the curvature hypothesis is to guarantee that the family $\{(M,\ \lambda^{-1}g) : \lambda \geq 1\}$ has curvature bounded below λ, away from the base point. (In this set-up, the base point p_0 in general becomes a singular point in a limit as $\lambda_i \to +\infty$, corresponding to the curvature at p_0 "blowing up" in the rescaling.) This lower bound on curvature arises from the fact that distance is rescaling by $1/\lambda$ and the curvature by λ^2. So a lower bound on $\liminf r^2 K(\sigma)$ as $r \to +\infty$ gives a lower curvature bound at points fixed distance from p_0 that is uniform over variation of λ, for $r > 0$ fixed.

According to [Grove and Petersen 1991], a space arising as a GH limit of Riemannian manifolds (of fixed dimension) with curvature bounded uniformly below has constant dimension: The GH limit of a sequence of such manifolds of dimension n is either again of dimension n, with some metric singularities perhaps, or alternatively "collapsing" occurs everywhere, and the limit space has everywhere a fixed dimension $k < n$. In either case the limit space is an Alexandrov space. This result applies, in particular, to the GH limit of a GH-convergent sequence $\{(M,\ \lambda_i^{-1}g)\}$, where $\lim \lambda_i = +\infty$. The collapsing case (into a nonmanifold) can indeed occur, as one sees for instance from a capped-off two-dimensional half-infinite cylinder, with curvature ≥ 0 everywhere and curvature $\neq 0$ outside a compact set: the limit space in this case is isometric to $\{t \in \mathbb{R} : t \geq 0\}$.

In the case of GH limits of sequences $\{(M_i, g_i)\}$ of manifolds with curvature bounded both above and below, stronger conclusions hold than in the more general situation of [Grove and Petersen 1991], where the curvature is bounded below only. Collapsing can still occur, as the example of Berger spheres shows.

But with bounds both above and below and with *no* collapsing, the limit space is (identifiable with) an n-dimensional Riemannian manifold with a $C^{1,\alpha}$ metric.

In this no collapsing situation, [Grove and Petersen 1991] guarantees that injectivity radius is bounded away from zero uniformly (or uniformly on compact sets, in case the (M_i, g_i) are noncompact, pointed). Thus the $C^{1,\alpha}$ statement is just a restatement of the standard $C^{1,\alpha}$ Convergence Theorem proposed by Gromov [1981b] and established in detail in [Greene and Wu 1988], and independently in [Peters 1987], using results from [Jost and Karcher 1982] (compare [Nikolaev 1980]).

This line of reasoning applies in particular to a family of the form $\{(M,\, \lambda^{-1}g) : \lambda \geq 1\}$ if M is a manifold of faster than quadratic curvature decay. The family has (sectional) curvature bounded both below and above (away from the point p_0, that is, on the complement in each $(M,\, \lambda^{-1}g)$ of a ball in that $(M,\, \lambda^{-1}g)$ around p_0 of any fixed positive radius $\varepsilon > 0$, the curvature bounds depending on ε). This follows by the reasoning of the first paragraph of this section.

Suppose that $\{(M,\, \lambda_i^{-1}g)\}$, with $\lim \lambda_i = +\infty$, is a GH convergent sequence in this setting and that the convergence occurs without collapsing. (The occurrence of the single singular point p_0 at which curvature is possibly unbounded causes no difficulty in the applications of [Grove and Petersen 1991] and the $C^{1,\alpha}$ convergence theorem that is used here.) Then the limit space (with p_0 ignored) is a Riemannian manifold with $C^{1,\alpha}$ Riemannian metric, and this metric is flat, in the sense for instance that it is the limit of C^∞ metrics with curvature going to zero, so that it is also flat in the triangle comparison sense. Thus the limit must be in fact a C^∞ Riemannian manifold of zero curvature. If M is connected at infinity and simply connected at infinity, this limit flat manifold with the isolated singularity p_0 removed is isometric to $\mathbb{R}^n$ with a single point removed (compare Theorem 6.1). Note that the metric space structure limit includes p_0 so that p_0 is metrically isolated too, that is the limit space minus p_0 is isometric to $\mathbb{R}^n$ minus a point, not to $\mathbb{R}^n$ minus a set with more than one point). Since $n \geq 3$, the limit space is in fact simply isometric to $\mathbb{R}^n$.

In this situation, one concludes that M is "euclidean at infinity", not only in the sense that M minus some compact set is topologically $\mathbb{R}^n$ minus some compact set, but also that the metric structure of M converges (in $C^{1,\alpha}$) to that of $\mathbb{R}^n$ in the following sense: There are coordinates on M that amount to a map of $\mathbb{R}^n$ minus a (large) ball into M in which coordinates the metric g converges to the Euclidean $g_{ij} = \delta_{ij}$ metric with increasing distance from p_0. This program is carried out in detail in [Bando et al. 1989].

Continuing in this faster-than-quadratic decay setting, one notes that the required "no collapsing" can be guaranteed by assuming Euclidean volume growth in the sense already discussed, that is, that $r^{-n} \operatorname{vol} B(p_0, r)$ bounded away from 0 as $r \to +\infty$. In this set-up, the assumption of simple connectivity made earlier for convenience of exposition is not in fact needed: The limit flat Riemannian manifold with an isolated singular point p_0 must have finite fundamental group (otherwise it could not have euclidean volume growth, as one sees from the metric classification of manifolds flat outside a compact set in [Schroeder and Ziller 1989] or [Greene and Wu 1993]). So by passing to a finite cover, one can reason as in the simply connected case to conclude that M "at infinity" converges to $(S^{n-1}/\Gamma^1) \times \mathbb{R}^+$, where Γ is a finite group of fixed-point-free isometries of S^{n-1}: see [Bando et al. 1989] for details.

This perspective explains the repeated occurrence of volume estimation as the vital point in the proofs of the "gap theorems" of the previous section: euclidean volume growth is the natural guarantee of "good", that is, essentially euclidean, behavior at infinity.

The convergence viewpoints so far discussed in relation to the gap theorems and asymptotically locally euclideanness are also related to the concept of ideal boundaries for open manifolds defined in terms of geodesic geometry. This concept was first introduced in substantial form by Eberlein and O'Neill [1973] for manifolds of nonpositive curvature. Two rays emanating from a base point p_0 were to be regarded as equivalent if they remained a bounded distance apart; and the ideal boundary, what is now known as the (boundary of) the Eberlein–O'Neill compactification, was obtained by introducing an appropriate metric on the equivalence classes of rays.

For manifolds of nonnegative curvature, a corresponding idea was suggested in [Ballmann et al. 1985] with some "exercises" pointing the way to detailed development. Further development in detail indeed was carried out by Kasue [1988], Shioya [1988], and Shiohama [a]. In summary form, the ideal boundary or "space at infinity" consists of equivalence classes of rays emanating from a base point p_0. Two rays γ_1 and γ_2 are taken to be equivalent if $\lim t^{-1} \operatorname{dist}(\gamma_1(t), \gamma_2(t)) = 0$ as $t \to +\infty$.

Then the distance between two equivalence classes, represented by, say, $\alpha(t)$ and $\beta(t)$, is defined to be $\lim t^{-1} \operatorname{dist}_t(A_t, B_t)$, where A_t and B_t are the intersections of the rays α and β with the radius-t sphere around p_0, and dist_t is the distance obtained by using the intrinsic metric of the radius-t sphere around p_0. (Of course, this intrinsic metric distance differs considerably, in general, from the distance in M between the two points.)

For manifolds of nonnegative curvature, the space at infinity or ideal boundary in this geodesic sense is related to the GH limiting ideas as follows [Shiohama a]: If (M, g) is a complete noncompact manifold of nonnegative sectional curvature, with fixed base point p_0, the GH limit of the family $(M, \lambda^{-1}g)$ as $\lambda \to +\infty$ exists and is a cone over a compact (metric) space $M(\infty)$. The space $M(\infty)$ is isometric to the space at infinity in the geodesic sense.

At first sight, the space $M(\infty)$ might seem to be quite arbitrary. It has the property that it is an Alexandrov space of curvature ≥ 1 (see [Guijarro and Kapovitch 1995] for history), but no other restrictions spring immediately to mind. Thus the following result of Luis Guijarro and Vitali Kapovitch [1995] is rather surprising:

THEOREM 7.1. *If (M, g) is a complete, noncompact, n-dimensional manifold of nonnegative curvature and if $M(\infty)$ is connected, there is a locally trivial fibration $f : S^k \to M(\infty)$ for some $k \leq n - 1$.*

From this line of thought, one can deduce that there are compact manifolds of curvature ≥ 1 that do not occur as the ideal boundary $M(\infty)$ of any complete noncompact M of nonnegative curvature.

References

[Abresch 1985] U. Abresch, "Lower curvature bounds, Toponogov's theorem, and bounded topology I", *Ann. scient. Éc. Norm. Sup.* (4) **18** (1985), 651–670.

[Abresch and Gromoll 1990] U. Abresch and D. Gromoll, "On complete manifolds with nonnegative Ricci curvature", *J. Amer. Math. Soc.* **3** (1990), 355–374.

[Allendoerfer and Weil 1943] C. Allendoerfer and A. Weil, "The Gauss–Bonnet theorem for Riemannian polyhedra", *Trans. Amer. Math. Soc.* **53** (1943), 101–129.

[Anderson and Cheeger 1992] M. Anderson and J. Cheeger, "C^α compactness for manifolds with Ricci curvature and injectivity radius bounded from below", *J. Diff. Geom.* **35** (1992), 265–281.

[Bando et al. 1989] S. Bando, A. Kasue, and H. Nakajima, "A construction of coordinates at infinity on manifolds with fast curvature decay and maximal volume growth", *Invent. Math.* **97** (1989), 313–349.

[Ballmann et al. 1985] W. Ballmann, M. Gromov, and V. Schroeder, *Manifolds of Nonpositive Curvature*, Birkhäuser, Boston, 1985.

[Becker et al. 1980] J. Becker, C. W. Henson, and L. Rubel, "First-order conformal invariants", *Ann. of Math.* **112** (1980), 123–178.

[Bishop and Crittenden 1964] R. Bishop and R. Crittenden, *Geometry of Manifolds*, Academic Press, New York, 1964.

[Burago et al. 1992] Y. Burago, M. Gromov, and G. Perelman, "A. D. Aleksandrov spaces with curvatures bounded below", *Uspehi Mat. Nauk* **47**:2 (1992), 3–51, 222; translation in *Russian Math. Surveys* **47**:2 (1992), 1–58.

[Busemann 1958] H. Busemann, *Convex Surfaces*, Tracts in Pure and Appl. Math. **6**, Interscience, New York, 1958.

[Calabi 1957] E. Calabi, "An extension of E. Hopf's maximum principal with an application to geometry", *Duke Math. J.* **25** (1957), 45–56.

[Cai 1991] M. Cai, "Ends of Riemannian manifolds with nonnegative Ricci curvature outside a compact set", *Bull. Amer. Math. Soc.* **24** (1991), 371–377.

[Cai et al. 1994] M. Cai, G. Galloway and Z. Liu, "Local splitting theorems for Riemannian manifolds", *Proc. Amer. Math. Soc.* **120** (1994), 1231–1239.

[Cheeger and Gromoll 1971] J. Cheeger and D. Gromoll, "The splitting theorem for manifolds of nonnegative Ricci curvature", *J. Differential Geom.* **6** (1971), 119–129.

[Cheeger and Gromoll 1972] J. Cheeger and D. Gromoll, "On the structure of complete manifolds with nonnegative curvature", *Ann. of Math.* **96** (1972), 413–443.

[Cheeger and Tian 1994] J. Cheeger and G. Tian, "On the cone structure at infinity of Ricci flat manifolds with Euclidean volume growth and quadratic curvature decay", *Invent. Math.* **118** (1994), 493–571.

[Cohn-Vossen 1935] S. Cohn-Vossen, "Kurzeste Wege und Totalkrümmung auf Flächen", *Compositio Math.* **2** (1935), 69–133.

[Drees 1994] G. Drees, "Asymptotically flat manifolds of nonnegative curvature", *Differential Geom. Appl.* **4** (1994), 77–90.

[Eberlein and O'Neill 1973] P. Eberlein and B. O'Neill, "Visibility manifolds", *Pacific J. Math.* **46** (1973), 45–109.

[Elencwajg 1975] G. Elencwajg, "Pseudoconvexité locale dans les variétés kahlériennes, *Ann. Inst. Fourier* **25** (1975), 295–314.

[Elerath 1979] D. Elerath, "Open nonnegatively curved 3-manifolds with a point of positive curvature", *Proc. Amer. Math. Soc.* **75** (1979), 92–94.

[Eschenburg et al. 1989] J. Eschenburg, V. Schroeder, and M. Strake, "Curvature at infinity of open nonnegatively curved manifolds", *J. Differential Geom.* **30** (1989), 155–166.

[Gao 1990] L. Gao, "$L^{n/2}$ curvature pinching", *J. Diff. Geo.* **32** (1990), 713–774.

[Greene 1989] R. E. Greene, "Some recent developments in Riemannian geometry", pp. 1–30 in *Recent Developments in Riemannian Geometry* (Los Angeles, 1987), edited by S.-Y. Cheng et al., Contemp. Math. **101**, Amer. Math. Soc., Providence, 1989.

[Greene and Shiohama 1981a] R. E. Greene and K. Shiohama, "Convex functions on complete noncompact manifolds: topological structure", *Invent. Math.* **63** (1981), 129–157.

[Greene and Shiohama 1981b] R. E. Greene and K. Shiohama, "Convex functions on complete noncompact manifolds: differentiable structure", *Ann. scient. Éc. Norm. Sup.* (4) **14** (1981), 357–367.

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

[Greene and Wu 1974] R. E. Greene and H. Wu, "Integrals of subharmonic functions on manifolds of nonnegative curvature", *Invent. Math.* **27** (1974), 265–298.

[Greene and Wu 1976] R. E. Greene and H. Wu, "C^∞ Convex functions and manifolds of positive curvature", *Acta Math.* **137** (1976), 209–245.

[Greene and Wu 1977] R. E. Greene and H. Wu, "Analysis on noncompact Kähler manifolds", pp. 69–100 in *Several complex variables* (Williams College, 1975), edited by R. O. Wells, Jr., Proc. Sym. Pure Math. **30**, Amer. Math. Soc., Providence, 1977.

[Greene and Wu 1978] R. E. Greene and H. Wu, "On Kähler manifolds of positive bisectional curvature and a theorem of Hartogs", *Abh. Math. Sem. Univ. Hamburg* **47** (1978), 171–185.

[Greene and Wu 1979] R. E. Greene and H. Wu, *Function theory on manifolds which possess a pole*, Lecture Notes in Math. **699**, Springer, Berlin, 1979.

[Greene and Wu 1982] R. E. Greene and H. Wu, "Gap theorems for noncompact Riemannian manifolds", *Duke Math. J.* **49** (1982), 731–756.

[Greene and Wu 1988] R. E. Greene and H. Wu, "Lipschitz convergence of Riemannian manifolds", *Pacific J. Math.* **131** (1988), 119–141.

[Greene and Wu 1993] R. E. Greene and H. Wu, "Nonnegatively curved manifolds which are flat outside a compact set", pp. 327–335 in *Differential geometry: Riemannian geometry* (Los Angeles, 1990), edited by R. Greene and S. T. Yau, Proc. Symp. Pure Math. **54** (3), Amer. Math. Soc., Providence, 1993.

[Greene et al. 1994] R. E. Greene, P. Petersen, and S. H. Zhu, "Riemannian manifolds of faster-than-quadratic curvature decay", *Internat. Math. Res. Notices* **1994**, no. 9, 363–377.

[Gromoll and Meyer 1969] D. Gromoll and W. Meyer, "On complete manifolds of positive curvature", *Ann. of Math.* **90** (1969), 75–90.

[Gromov 1981a] M. Gromov, "Curvature, diameter and Betti numbers", *Comment. Math. Helv.* **56** (1981), 179–195.

[Gromov 1981b] M. Gromov, *Structures métriques pour les variétés riemanniennes*, edited by J. Lafontaine and P. Pansu, CEDIC, Paris, 1981.

[Gromov 1986] M. Gromov, *Partial Differential Relations*, Springer, Berlin, 1986.

[Gromov and Rohlin 1970] M. Gromov and V. Rohlin, "Imbeddings and immersions in Riemannian geometry", *Uspehi Mat. Nauk* **25**:5 (1970), 3–62; translation in *Russian Math. Surveys* **25**:5 (1970), 1–58.

[Grove and Petersen 1991] K. Grove and P. Petersen, "Manifolds near the boundary of existence", *J. Differential Geom.* **33** (1991), 379–394.

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

[Guijarro a] L. Guijarro, in preparation.

[Guijarro and Kapovitch 1995] L. Guijarro and V. Kapovitch, "Restrictions on the geometry at infinity of nonnegatively curved manifolds", *Duke Math. J.* **78** (1995), 257–276.

[Guijarro and Petersen a] L. Guijarro and P. Petersen, "Rigidity in nonnegative curvature", preprint.

[van Heijenoort 1952] J. van Heijenoort, "On locally convex manifolds", *Comm. Pure Appl. Math.* **5** (1952), 223–242.

[Hartmann and Nirenberg 1951] Hartmann and L. Nirenberg, "On the spherical image maps whose Jacobians do not change signs", *Amer. J. Math.* **73** (1951).

[Jost and Karcher 1982] J. Jost and H. Karcher, "Geometrische methoder zur Gewinnung von a-priori-Schranken für harmonische Abbildungen", *Manuscripta Math.* **40** (1982), 27–77.

[Karcher 1989] H. Karcher, "Riemannian comparison constructions", pp. 170–222 in *Global differential geometry*, edited by S.-S. Chern, MAA Stud. Math. **27**, Math. Assoc. America, Washington, DC, 1989.

[Kasue 1988] A. Kasue, "A compactification of a manifold with asymptotically nonnegative curvature", *Ann. scient. Éc. Norm. Sup.* (4) **21** (1988), 593–622.

[Kasue and Sugahara 1987] A. Kasue and K. Sugahara, "Gap theorems for certain submanifolds of euclidean spaces and hyperbolic space forms", *Osaka J. Math.* **24** (1987), 679–704.

[Li and Tam 1992] P. Li and L.-F. Tam, "Harmonic functions and the structure of complete manifolds", *J. Diff. Geom.* **35** (1992), 359–383.

[Lohkamp 1992] J. Lohkamp, "Negatively Ricci curved manifolds", *Bull. Amer. Math. Soc.* (2) **27** (1992), 288–291.

[Marenich 1981] V. B. Marenich, "Metric structure of open manifolds of nonnegative curvature" (Russian), *Dokl. Akad. Nauk SSSR* **261** (1981), 801–804.

[Mok et al. 1981] N. Mok, Y. T. Siu, and S. T. Yau, "The Poincaré–Lelong equation on complete Kähler manifolds", *Compositio Math.* **44** (1981), 183–218.

[Nikolaev 1980] I. G. Nikolaev, "Parallel translation and smoothness of the metric of spaces of bounded curvature", *Dokl. Akad. Nauk SSSR* **250** (1980), 1056–1058.

[Özaydin and Walschap 1994] M. Özaydin and G. Walschap, "Vector bundles with no soul", *Proc. Amer. Math. Soc.* **120** (1994), 565–567.

[Perelman 1994] G. Perelman, "Proof of the soul conjecture of Cheeger and Gromoll", *J. Differential Geom.* **40** (1994), 209–212.

[Peters 1987] S. Peters, "Convergence of Riemannian manifolds", *Compositio Math.* **62** (1987), 3–16.

[Poor 1974] W. A. Poor, "Some results on nonnegatively curved manifolds", *J. Differential Geom.* **9** (1974), 583–600.

[Rushing 1973] T. Rushing, *Topological embeddings*, Academic Press, New York, 1973.

[Sacksteder 1960] R. Sacksteder, "On hypersurfaces with no negative curvature", *Amer. J. Math.* **82** (1960), 609–630.

[Schroeder and Ziller 1989] V. Schroeder and W. Ziller, "Rigidity of convex domains in manifolds with nonnegative Ricci and sectional curvature", *Comment. Math. Helv.* **64** (1989), 173–186.

[Sharafutdinov 1977] V. Sharafutdinov, "Pogorelov–Klingenberg theorem for manifolds homeomorphic to $\mathbb{R}^n$", *Sibirsk. Math. Zh.* **18** (1977), 915–925.

[Shiohama a] K. Shiohama, "An introduction to the geometry of Alexandrov spaces", lecture notes, Seoul National University.

[Shioya 1988] T. Shioya, "Splitting theorems for nonnegatively curved manifolds", *Manuscripta Math.* **61** (1988), 315–325.

[Siu and Yau 1977] Y. T. Siu and S. T. Yau, "Complete Kähler manifolds of faster than quadratic curvature decay", *Ann. Math.* **105** (1977), 225–264.

[Strake 1988] M. Strake, "A splitting theorem for open nonnegatively curved manifolds", *Manuscripta Math.* **61** (1988), 315–325.

[Toponogov 1959] V. Toponogov, "Riemann spaces with curvature bounded below" (Russian), *Uspehi Mat. Nauk* **14** (1959), 85, 87–130.

[Toponogov 1964] V. Toponogov, "Spaces with straight lines", *Amer. Math. Soc. Transl.* **37** (1964), 287–290.

[Unnebrink a] S. Unnebrink, "Asymptotically flat 4-manifolds", preprint.

[Walschap 1988] G. Walschap, "Nonnegatively curved manifolds with souls of codimension 2", *J. Differential Geom.* **27** (1988), 525–537.

[Wu 1970] H. Wu, "An elementary method in the study of nonnegative curvature", *Acta Math.* **142** (1970), 57–78.

[Wu 1987] H. Wu, "On manifolds of partially positive curvature", *Indiana Univ. Math. J.* **36** (1987), 525–548.

[Yang 1992] D. Yang, "Convergence of Riemannian manifolds with integral bounds on curvature I–II", *Ann. Sci. École Norm. Sup.* **25** (1992), 77–105, 179–199.

[Yim 1988] J. W. Yim, "Distance nonincreasing retraction on a complete open manifold of nonnegative sectional curvature", *Ann. Global Anal. Geom.* **6** (1988), 191–206.

Robert E. Greene
Department of Mathematics
University of California
Los Angeles, CA 90024
greene@math.ucla.edu

Support, Comparison and Splitting

Second variation support functions (Calabi [1957] et al.)

Toponogov triangle comparison [1959]
(Second variation ⇒ Toponogov: H. Karcher [1989])

|

Toponogov splitting [1964]
(second curvature ≥ 0 and existence of a line ⇒ isometric to $N \times \mathbb{R}$)

|

Cheeger–Gromoll splitting [1971]
(Ricci ≥ 0 and existence of a line ⇒ isometric to $N \times \mathbb{R}$)

|

Greene and Wu [1974], Greene and Shiohama [1981a]
(sectional curvature ≥ 0 outside a compact ⇒ finite number of product ends)

|

P. Li and L. Tam [1992]
(Ricci curvature ≥ 0 outside a compact ⇒ finite number of ends)
(Geometric proof by M. Cai [1991])

|

Cai, Galloway, Liu [1994] (localized version)

Related support function applications

Elenczwag [1975], Greene and Wu [1978] (Kähler manifolds of nonnegative holomorphic bisectional curvature, plurisubharmonic functions, etc.)

H. Wu [1987] (mixed curvature conditions, q-convexity)

Convex Function Theory

Greene and Wu [1973; 1976]
(approximation of convex functions by almost convex C^∞ functions; existence of C^∞ strictly convex exhaustion function under positive curvature)

Sharafutdinov [1977] (distance nonincreasing retraction onto soul)

Greene and Shiohama [1981a; 1981b] (structure of convex functions, distance-nonincreasing retraction in general convex function case)

Yim [1988] (space of souls)

G. Perelman [1994] (rigidity of retraction onto soul)

Open Manifolds of Nonnegative Curvature

Cohn-Vossen [1935] (surfaces: $\int k\,dA \le 2\pi\chi(M)$, classification for $k \ge 0$)

Toponogov triangle comparison [1959]

Gromoll and Meyer [1969] (curvature $> 0 \Rightarrow$ homeomorphic to $\mathbb{R}^n$)

Cheeger and Gromoll [1972]

Soul Theorem
(curvature $\ge 0 \Rightarrow$ homeomorphic to total space of bundle over compact manifold of curvature ≥ 0)
(classification of codimension-one soul case)

Conjecture
(curvature ≥ 0 everywhere and > 0 somewhere $\Rightarrow$ homeomorphic to $\mathbb{R}^n$)

Elerath [1979] (true if manifold not too curved)

G. Perelman [1994] (true)

Poor [1974] (diffeomorphism in all dimensions, generalizing Cohn-Vossen $\le$)

Marenich [1981], Strake [1988] (splitting when holonomy is trivial)

Abresch [1985] (asymptotically nonnegative $\Rightarrow$ estimatable finite topology)

Walschap [1988] (codimension-two soul)

Özaydin and Walschap [1994] (not all bundles over compact manifolds of curvature ≥ 0 occur)

Rescaling and Convergence

Jost and Karcher [1982] (harmonic coordinates)

Gromov [1981b] (Gromov–Hausdorff and other convergence ideas, rescaling)

Peters [1987], Greene and Wu [1988] ($C^{1,\alpha}$ convergence); cf. Nikolaev [1980]

Bando, Kasue, and Nakajima [1989] (curvature decay and volume growth of order $n \Rightarrow$ asymptotically locally euclidean at ∞)

Cheeger and Tian [1994] (cone structure for Kähler case)

Anderson and Cheeger [1992], Gao [1990], Yang et al. [1992] (integral curv., Ricci)

Behavior at Infinity: Gap Theorems

Greene and Wu [1977] (conjecture on Kähler manifolds: $\pi_1 = 0$, curvature ≤ 0 and decaying faster than quadratically $\Rightarrow$ biholomorphic to $\mathbb{C}^n$)

Siu and Yau [1977] (proof of conjecture)

Mok, Siu, and Yau [1981] (same hypotheses $\Rightarrow$ isometric to $\mathbb{C}^n$ if $n \geq 2$)

Greene and Wu [1982] (Riemannian gap theorems:
$\pi_1 = 0$ at ∞, flat outside a compact, curvature of one sign $\Rightarrow$ flat;
curvature of one sign,
decay faster than quadratic, pole, dim $\neq 4, 8$ if curvature $\geq 0 \Rightarrow$ flat)

Kasue and Sugahara [1987] (pole case, $n = 4, 8$)

Eschenburg, Schroeder, and Strake [1989]
(codimension of soul ≥ 3, curvature $\to 0 \Rightarrow$ soul flat)
(curvature decay rates for gap theorems in odd dimensions in terms of codimension of soul)

Drees [1994] (curvature ≥ 0, dim $\neq 4, 8$, decay faster than quadratic $\Rightarrow$ flat)

Greene, Petersen, and Zhu [1994]
(π_1 finite at ∞, decay faster than quadratic (if dim $\neq 4, 8$) or quartic (else) $\Rightarrow$ quotient of asymptotically euclidean manifold)
(π_1 infinite, faster than quartic $\Rightarrow$ equivalent metric flat outside a compact)

Guijarro and Petersen [a] (curvature $\to 0$ at $\infty \Rightarrow$ soul flat)

Schroeder and Ziller [1989]
(classification of flat outside compact set; by another method, Greene and Wu [1993])

Unnebrink [a] (counterexample to faster than quadratic growth in dimension 4)