MANIFOLDS OF NEGATIVE CURVATURE

BY

R. L. BISHOP AND B. O'NEILL[1]

1. **Introduction**   A C^∞ function f on a riemannian manifold M is *convex* provided its hessian (second covariant differential) is positive semidefinite, or equivalently if $(f \circ \sigma)'' \geq 0$ for every geodesic σ in M. We shall apply this notion in a variety of ways to the study of manifolds of negative or nonpositive curvature.

Convexity has, of course, long been associated with negative curvature, but convex functions seem to have been used only locally or along curves. In the first part of this paper we give an abstract global treatment. Nonconstant convex functions exist only on manifolds of infinite volume (2.2); the first question about such a function on M (complete, $K \leq 0$) is whether it has a critical point—necessarily an absolute minimum. If not, M is diffeomorphic to a product $L \times R^1$ (3.12). If so, much of the topology and geometry of M is determined by the minimum set C of f. This comes about as follows. Like any set $\{m \in M \mid f(m) \leq a\}$, C is *totally convex*, that is, contains a geodesic segment σ whenever it contains the endpoints of σ. Let A be an arbitrary closed, totally convex set in M. In case A is a submanifold, it is totally geodesic and M is, via exponentiation, its normal bundle (3.1). This situation does not change greatly if A is not a submanifold (e.g., 3.4); A is always a topological manifold with boundary (possibly nonsmooth), whose interior is a locally totally geodesic submanifold. We describe a number of geometrically significant ways of constructing convex functions (4.1, 4.2, 4.8, 5.5, etc.); these show in particular that C may or may not be a submanifold.

In the second part of the paper we define and study the *mobility sequence* of a nonpositive curvature manifold M. The basic fact is that the set $P(M)$ of common zeroes of all Killing fields on M is a closed, totally convex submanifold (5.1). Thus M is a vector bundle over $P(M)$, which is totally geodesic and hence again has $K \leq 0$. The mobility sequence is then constructed by iteration: $M \to P(M) \to \cdots \to P^k(M) = Q$. It terminates with a submanifold that is either *mobile* ($P(Q)$ empty) or *immobile* ($P(Q) = Q$). We prove that if Q is mobile, or if $\pi_1(M)$ has nontrivial center, then (with a trivial exception) M is a product $L \times R^1$ and if also M contains a closed geodesic then in particular M is a vector bundle over a circle (4.9, 6.4). Since $P(M)$ is invariant under all isometries of M, the mobility sequence is closely related to the isometry group of M (e.g. 8.1).

We introduce the notion of *warped product* (or, more generally, *warped bundle*),

---

which uses convex functions to construct a wide variety of manifolds of negative curvature (7.5). For example, negative space forms can easily be constructed in this way from flat space forms (7.10). The study of Killing fields on warped products (8.2) shows in particular that the dimensions and length of mobility sequences are restricted by only the most trivial necessary conditions (8.4).

In the final part of the paper we investigate the fundamental group and for the principal results require strictly negative curvature, $K \leq c < 0$. We begin by relating convex functions to asymptotic geodesics, showing for example that if M admits a strictly convex function without minimum (and is not simply connected) then there is a unique equivalence class of asymptotic geodesics with properties decisive for the geometry of M (9.12–9.14). Our main tool in the study of the fundamental group is the notion of *ray group*: fix a point $m \in M$, then associated with each geodesic ray ρ starting at m is a subgroup of $\pi_1(M)$ consisting of those elements that (roughly speaking) decrease in length as they move out ρ. There are two types of such ray groups, depending on whether the ray diverges or has a limit cycle (10.10)—the latter type are always infinite cyclic. The fundamental group is the disjoint union of its ray groups (10.11) and these have the algebraic property that "nontrivial conjugates are disjoint" (10.12). In the compact case, for example, our results reduce to a sharpened form (10.16) of the Preissmann result that commutative subgroups are cyclic.

We assume that riemannian manifolds and the apparatus on them are C^∞ and that manifolds are connected—unless these points are raised. By a *closed geodesic* γ in M we mean a nonconstant periodic geodesic $\gamma : R \to M$. The closed interval $[0, 1]$ is denoted by I.

2. Convex functions. If f is a C^∞ real-valued function on a riemannian manifold M, its hessian $\nabla^2 f = \nabla \, df$ is the $(0, 2)$ tensor field such that

$$\nabla^2 f(X, Y) = XYf - (\nabla_X Y)f$$

for all vector fields X, Y on M. At each point $m \in M$, $\nabla^2 f$ is a symmetric bilinear form on the tangent space M_m. If γ is a geodesic and γ' denotes its tangent field, $\nabla^2 f(\gamma', \gamma') = (f \circ \gamma)''$.

We denote the covariant differential of a vector field X by A_X. Thus A_X is the $(1, 1)$ tensor field for which the value on a vector field Y is $A_X Y = \nabla_Y X$. The following properties of A_X are well known. The symmetry of A_X with respect to the riemannian inner product is equivalent to X being locally a gradient. If $X = \text{grad} \, f$, then A_X is the symmetric transformation given by $\nabla^2 f$ and the metric: $\nabla^2 f(Y, Z) = \langle A_X Y, Z \rangle$. The skew-symmetry of A_X is equivalent to X being a Killing field. The trace of A_X is the divergence δX of X, that is, the derivative of the volume expansion given by the flow transformations of X.

We say that a C^∞ function f is *strictly convex* if, at each pont, $\nabla^2 f$ is positive definite, and f is *convex* if, at each point, $\nabla^2 f$ is positive semidefinite. As we shall see, convex functions arise naturally on manifolds of nonpositive curvature.

2.1 PROPOSITION. *Let f be a convex function on a complete riemannian manifold M. Then*

(1) *The critical points of f are its absolute minimum points.*

(2) *For each number c, the set $M^c = \{m \in M \mid f(m) \leqq c\}$ is totally convex, that is, contains every geodesic segment of M whose endpoints are in M^c. In particular, M^c is connected.*

(3) *f is constant on each closed geodesic in M.*

(4) *For each value c of f, the inclusion $M^c \subset M$ induces a homomorphism of the fundamental group $\pi_1(M^c)$ onto $\pi_1(M)$.*

Proof. (1) It suffices to show that if p is a critical point, it is an (absolute) minimum point. If $m \in M$, let $\tau : I \to M$ be a geodesic segment from p to m, and let $g = f \circ \tau$. Then $g'(0) = 0$, since p is a critical point, and $g'' \geqq 0$, since f is convex. Hence $f(m) - f(p) = g(1) - g(0) \geqq 0$.

(2) If $\tau : I \to M$ is a geodesic segment with $f\tau(0) \leqq c$ and $f\tau(1) \leqq c$, then $(f\tau)'' \geqq 0$ implies $f\tau \leqq c$.

(3) If γ is a closed geodesic, then by (2), $f\gamma(s) \leqq f\gamma(t)$ for all s, $t \in R$. Thus $f\gamma$ is constant.

(4) Each element of $\pi_1(M, m)$, $m \in M^c$, can be represented by a geodesic loop τ at m. But by (2), τ lies in M^c.

The notion of total convexity obviously differs in an essential way from the usual notion of convexity in a riemannian manifold [1], [2]. For example, in a hyperboloid of revolution the minimal circle is totally convex, but a single point is not.

We call attention to the following papers, both appearing contemporaneously in the Annals of Mathematics: D. Gromoll and W. Meyer, *Manifolds of positive curvature*; J. Cheeger and D. Gromoll, *On the structure of complete manifolds of nonnegative curvature*. The notion of totally convex sets is used in both of these. In the second, it is shown that the inclusion $M^c \subset M$ above induces a homotopy equivalence; more generally, the same result obtains for a closed totally convex set C such that there are two points p, q in the interior of C for which all geodesic segments from p to q also lie in the interior of C. This condition is clearly satisfied by M^c. Some of our results below are immediate corollaries; for example, Proposition 2.5 (2). However, our proofs are more elementary, since in particular we do not use any Morse theory.

Clearly any intersection of totally convex sets is totally convex, and by the argument for (4) above the inclusion map of a totally convex subset into M induces a homomorphism onto of fundamental groups. Except in the case of nonpositive curvature, the proof that this inclusion induces an isomorphism seems to require a Morse theory argument.

According to the following proposition, many riemannian manifolds, in particular the compact ones, do not admit nontrivial convex functions.

2.2 PROPOSITION. *If M is a complete riemannian manifold having finite volume, then all convex functions on M are constant.*

Proof. Suppose f is a nonconstant convex function. Let $G = \operatorname{grad} f$ and let c be a noncritical value of f. Then $^cM = \{m \in M \mid f(m) \geqq c\}$ is a submanifold with boundary $f^{-1}(c)$ and nonzero volume, on which G is never zero and points inward from the boundary. Thus $G/\|G\| = U$ is a unit-length field defined on all of ^cM and "complete upwards" in the sense that the flow transformations $\{\phi_t\}$ of U are defined on all of ^cM for $t \geqq 0$. Since ϕ_t, $t > 0$, maps ^cM properly into a set of smaller volume, we must have that the divergence of U is negative somewhere on ^cM. However, for a local orthonormal basis $E_1 = U, E_2, \ldots, E_d$, we have

$$\delta U = \sum \langle A_U E_i, E_i \rangle$$

$$= \sum \langle \nabla_{E_i} U, E_i \rangle$$

$$= \sum \left(\langle G, E_i \rangle E_i \frac{1}{\|G\|} + \frac{1}{\|G\|} \langle A_G E_i, E_i \rangle \right)$$

$$= \|G\| E_1 \|G\|^{-1} + \|G\|^{-1} \sum \nabla^2 f(E_i, E_i).$$

But

$$E_1 \|G\|^{-1} = -\langle \nabla_U G, G \rangle \|G\|^{-3}$$

$$= -\langle A_G U, U \rangle \|G\|^{-2}$$

$$= -\nabla^2 f(U, U) \|G\|^{-2}.$$

This leaves $\delta U = \|G\|^{-1} \sum_{i=2}^d \nabla^2 f(E_i, E_i) \geqq 0$, a contradiction.

Fix the notation $C = C(f)$ for the set of critical points of a convex function f. By Proposition 2.1 (1), C is the minimum set of f. Thus if c is a nonminimum value of f, then M^c is a manifold with boundary the level hypersurface $f^{-1}(c)$. C itself (if nonempty) is not necessarily a manifold—with or without boundary. For example, if f is a convex function on R^1 with minimum set a closed interval, then on R^2 the function $f(x) + f(y)$ is convex with minimum set a square. The following lemma is an easy consequence of the computations $Uf = \|G\| > 0$ and $U^2 f = \langle \nabla_G G, G \rangle \|G\|^{-2} = \nabla^2 f(U, U) \geqq 0$.

2.3 LEMMA. *Let f be convex on a complete riemannian manifold M. The normalized gradient U (as in 2.2) is defined on M − C. If m ∉ C, the integral curve α of U such that α(0) = m is defined for all $t \geqq 0$, and fα is strictly increasing and convex. In particular fα(t) → ∞ as t → ∞.*

2.4 PROPOSITION. *Let f be a convex function without minimum on a complete riemannian manifold, and let $L = f^{-1}(c)$ be a level hypersurface of f. Then the restriction to $L \times (-\infty, 0]$ of the flow of $U = \operatorname{grad} f / \|\operatorname{grad} f\|$ is a diffeomorphism onto M^c.*

The proof is a straightforward consequence of the preceding lemma. It follows that ^cM is a strong deformation retract of M, and that $f^{-1}(c)$ is a strong deformation retract of M^c. In particular, (still assuming f has no minimum) *every level*

hypersurface of f is connected, since M^c is connected. This is not necessarily the case for a convex function with minimum, as is seen most simply by x^2 on R^1.

We now consider the strictly convex case.

2.5 PROPOSITION. *Let M be a complete riemannian manifold.* (1) *If M admits a strictly convex function, then M contains no closed geodesics.* (2) *If M admits a strictly convex function with a minimum, then M is contractible.*

Proof. (1) Closed geodesics are excluded by Proposition 2.1 (3), since $(f\gamma)''=0$ contradicts strict convexity.

(2) Suppose f is a strictly convex function with nonempty minimum set C. The hessian of f at each point of C is $\nabla^2 f$, which is nondegenerate, so the points of C are isolated. But C is connected by Proposition 2.1 (2), so C is a single point. The following lemma completes the proof.

2.6 LEMMA. *Let M be a complete riemannian manifold and let f be a convex function on M such that the minimum set C of f is nonempty and compact. Then C is a strong deformation retract of M.*

Proof. Let B_r be the compact neighborhood of C consisting of points in M at distance not more than r from C, where $r>0$.

We can choose r so small that for each $m \in B_r$ there is a unique geodesic segment which minimizes distance from m to C. For this we need only take $r<1$ so that every closed ball of radius $\leq r$ and center at some point in B_1 is (strictly) convex. Suppose $m \in B_r$ and that there were two minimizing geodesic segments α and β from C to m. The ends of α and β which lie in C, $\alpha(0)$ and $\beta(0)$, are the ends of a unique minimizing geodesic segment τ. Since the isosceles triangle $\alpha\beta\tau$ is contained in the convex ball with center m and radius $|\alpha|$, there are points on the base τ which are at distance less than $|\alpha|$ from m. But by Proposition 2.1 (2) τ lies in C, contradicting the minimality of α.

It now follows that C is a strong deformation retract of such a B_r. We simply deform B_r into C by moving points uniformly along the unique minimizing segments. Fix r at such a value henceforth.

Let c be the minimum value of f on the boundary ∂B_r of B_r. We claim $M^c \subset B_r$, so in particular M^c is compact. Indeed, if $f(m) \leq c$ and τ is a minimizing geodesic from C to m, then f is strictly increasing on τ (by convexity of f and the fact that only the point of τ in C is the initial point) and m is the maximum point of f on τ. It follows that no interior point of τ can be on ∂B_r and hence all of τ lies in B_r. Note that ∂B_r does not intersect C, so c is not a critical value of f.

Let U be the normalized gradient of f, defined on $M-C$. Then $Uf = \|\operatorname{grad} f\|$ has a minimum value $\mu>0$ on the compact set $f^{-1}(c)$. Now if $m \notin B_s$, where $s>r$, let τ be a minimal unit speed geodesic from M^c to m. Since $M^c \subset B_r$, $|\tau|>s-r$. We have that $\tau'(0)$ is orthogonal to $f^{-1}(c)$, so $\tau'(0) = U(\tau(0))$. Letting $g=f\tau$, this shows $g'(0)=\mu$ and hence $g' \geq \mu$. Thus $g(|\tau|)=f(m)>\mu(s-r)+c$. We have proved

that $M^b \subset B_s$, where $b = \mu(s-r) + c$, which shows that M^b is compact for every b.

We can now deform M into M^c by moving points backwards along the integral curves of U. This is possible because for any $m \in {}^c M$, the function Uf has a positive lower bound on $M^{f(m)} - M^c$, so the integral curve must eventually reach M^c.

2.7 REMARKS. By the same technique we can prove that if f is a convex function such that $f^{-1}(c)$ is compact for some c, then

(1) For any two values b, $b' > c$ of f the level hypersurfaces $f^{-1}(b)$ and $f^{-1}(b')$ are compact and diffeomorphic under the flow transformation $\phi_{b'-b}$ of the vector field grad $f/\|\text{grad } f\|^2$.

(2) If f has no minimum, M is diffeomorphic to $f^{-1}(c) \times R^1$. In fact, the flow of grad $f/\|\text{grad } f\|^2$ is a diffeomorphism from $f^{-1}(c) \times [0, \infty)$ onto ${}^c M$.

3. Convexity and nonpositive curvature. The results of the previous section can be strengthened when the manifold M has nonpositive curvature. For example, 2.7 (2) is obtained without the requirement that a level hypersurface be compact.

3.1 LEMMA. *A submanifold S of M (complete, $K \leq 0$) is closed and totally convex if and only if S is totally geodesic and the exponential map* exp: $\perp S \to M$ *is a diffeomorphism. ($\perp S$ denotes the normal bundle of S.)*

Proof. (1) Suppose S is a closed totally convex submanifold of M. Let $\alpha: I \to S$ be the unique shortest geodesic segment in S joining two points p, q which lie in some convex neighborhood of S. To show that S is totally geodesic it suffices to prove that α is a geodesic of M. Let $\beta: I \to M$ be a minimal geodesic segment in M from p to q. Since S is totally convex, β lies in S. Since $|\beta| \leq |\alpha|$ the uniqueness of α implies $\beta = \alpha$. Total convexity also implies that the inclusion $i: S \subset M$ induces a homomorphism $i_\#$ of $\pi_1(S)$ *onto* $\pi_1(M)$. Since S is closed and totally geodesic, the Hadamard-Herman theorem [4] asserts that exp: $\perp S \to M$ is a diffeomorphism. (Thus $i_\#$ is an isomorphism.)

(2) Suppose S is a totally geodesic submanifold of M such that exp: $\perp S \to M$ is a diffeomorphism. The latter condition implies that S is closed in M. Let $\tau: I \to M$ be a geodesic segment in M joining points $p, q \in S$. Since exp is a diffeomorphism, τ is fixed-endpoint homotopic to a curve in S from p to q, hence to a geodesic γ of S from p to q. But S is totally geodesic, so γ is a geodesic of M. Since $K \leq 0$, we have $\tau = \gamma$, hence τ lies in S.

Our aim now is to show that the content of the preceding lemma carries over, to some extent, to arbitrary closed totally convex sets (e.g. the minimum set of a convex function).

Let A be a totally convex subset of a complete manifold M. If $p \in A$, then $x \in M_p$ is *tangent to* A provided x is the initial velocity of some curve α that is initially in A (that is, $x = \alpha'(0)$ where $\alpha([0, \varepsilon]) \subset A$). If $m \in M$, a *perpendicular* from m to A is a geodesic segment $\tau: I \to M$ such that (1) $\tau(0) = m$, (2) $\tau(1) \in A$, and (3) $\langle \tau'(1), x \rangle \geq 0$ for every tangent x to A at $\tau(1)$. Note that if $m \in A$, then the

constant geodesic segment at m is the unique perpendicular from m to A. In fact, if τ is a geodesic loop at $m \in A$, then τ lies in A, so $\tau(1-t)$ is initially in A, whence $0 \leq \langle \tau'(1), -\tau'(1) \rangle$ and τ is constant.

It is easy to see that the set of tangents to A at p is a convex cone in M_p. Indeed, suppose $x, y \in M_p$ are tangents to A. It is already clear that tx, ty are tangent to A for every $t \geq 0$. Suppose that α, β are curves in A such that $\alpha'(0) = x$, $\beta'(0) = y$. For sufficiently small $t > 0$ there is a unique minimal geodesic segment γ_t from $\alpha(t) = \gamma_t(0)$ to $\beta(t) = \gamma_t(1)$. Since A is totally convex, γ_t lies in A. Then for $0 \leq s \leq 1$, $\tau(t) = \gamma_t(s)$ defines a curve in A such that $\tau'(0) = (1-s)x + sy$. An appropriate name for this property of A would seem to be *infinitesimally convex*. This notion occurs in control theory as the key property of the accessible set in the Pontryagin maximum principle; only the manifold structure, not the riemannian structure, is involved. Now it can be shown as in the paper of Cheeger and Gromoll (*op. cit.*) that a closed totally convex set is a locally totally geodesic submanifold with possibly nonsmooth boundary.

3.2 LEMMA. *Let M be a complete simply connected manifold with $K \leq 0$. Let A be a closed totally convex subset of M. Then*

(1) *For each $m \in M$ there exists a unique perpendicular τ_m from m to A.*

(2) *τ_m is the unique shortest geodesic segment from m to A.*

(3) *The function $\rho: M \to A$ sending m to $\tau_m(1)$ is a continuous retraction of M onto A.*

(4) *ρ is distance-nonincreasing (if $K < 0$, distance-decreasing except for points of A).*

Proof. If $m \in M$, then since A is closed, there is a point p of A nearest to m. Let τ_m be the geodesic segment from m to p. Since τ_m is a shortest geodesic from m to A, a standard argument shows that τ_m is a perpendicular to A. By a remark above we can assume $m \notin A$. Let α be a perpendicular from m to A and assume $\alpha \neq \tau_m$. The hypotheses on M show that $\tau_m(1) \neq \alpha(1)$; let γ be the geodesic segment from $\tau_m(1)$ to $\alpha(1)$. By total convexity, γ lies in A. It follows easily that the angle between γ and τ_m at $\tau_m(1)$ is at least $\pi/2$, and similarly for the angle between γ and α at $\alpha(1)$. Thus the sum of the angles of the geodesic triangle $\gamma \alpha \tau_m$ exceeds π. This contradiction shows that $\alpha = \tau_m$, proving (1) and (2). The continuity of ρ (in (3)) follows from (2). To prove (4), if $m, n \in M$ let $r: I \times I \to M$ be the rectangle such that $r(\cdot, v)$ is the geodesic segment from $\tau_m(v)$ to $\tau_n(v)$. Let $L(v)$ be the length of $r(\cdot, v)$. A computation gives the first variation of L^2 at $v = 1$ in terms of the inner products used in defining perpendicularity:

$$L^{2\prime}(1) = 2(-\langle \tau_n'(1), -\gamma'(1) \rangle - \langle \tau_m'(1), \gamma'(0) \rangle) \leq 0,$$

where $\gamma = r(\cdot, 1)$. Since $K \leq 0$, and τ_m and τ_n are geodesics, the second variation formula [1] reduces to

$$L^{2\prime\prime} = 2 \int_0^1 \left(\left\| \nabla_u \frac{\partial r}{\partial v} \right\|^2 - K\left(\frac{\partial r}{\partial u}, \frac{\partial r}{\partial v} \right) \left\| \frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v} \right\|^2 \right) du \geq 0.$$

Hence $L(1) \leq L(0)$; that is, $d(\rho(m), \rho(n)) \leq d(m, n)$. If $K < 0$ and one of m, n is not in A, then equality in the above formulas implies $\nabla_u(\partial r/\partial v) = 0$ and $\partial r/\partial u \wedge \partial r/\partial v = 0$, from which it easily follows that $m = n$.

3.3 REMARK. It may be of interest to examine the case of equality of distances, $d(\rho(m), \rho(n)) = d(m, n)$, in a little more detail. Clearly we may assume $\rho(m) \neq \rho(n)$ and one of m, n not in A. Since first variation vanishes, we must have that τ_m and τ_n are both perpendicular to the geodesic segment γ from $\rho(m)$ to $\rho(n)$. Since second variation vanishes, we have that $\nabla_u(\partial r/\partial v) = \nabla_v(\partial r/\partial u) = 0$, so the longitudinal and transverse fields of r are of constant length and perpendicular to each other. In particular, m and n are at the same distance from A. Moreover, the plane sections spanned by these fields have zero curvature. Consequently, the interior of r is a totally geodesic flat surface in M (but not complete).

We now eliminate the hypothesis of simple connectivity from the preceding lemma.

3.4 PROPOSITION. *Let A be a closed totally convex set in a complete riemannian manifold with $K \leq 0$. Then*

(1) *For each point $m \in M$ there is a unique perpendicular σ_m from m to A.*

(2) *σ_m is the unique shortest geodesic segment from m to A.*

(3) *The function $H(m, t) = \sigma_m(t)$, $(m, t) \in M \times I$, is a continuous deformation of M onto A.*

(4) *The continuous retraction $\rho: M \to A$ sending m to $\sigma_m(1)$ is distance-non-increasing (if $K < 0$, distance-decreasing except for points of A).*

Proof. Let $\pi: M_1 \to M$ be the simply connected riemannian covering of M. Clearly $A_1 = \pi^{-1}(A)$ is closed and totally convex. Let $\rho_1: M_1 \to A_1$ be the retraction obtained in the preceding lemma, and again, for $p \in M_1$, let τ_p be the geodesic segment from p to $\rho_1(p)$. If δ is a deck transformation, then, since $\delta(A_1) = A_1$, it follows that $\delta\rho_1(p)$ is the nearest point of A_1 to $\delta(p)$. Thus, by uniqueness of nearest points, $\delta\rho_1(p) = \rho_1\delta(p)$; that is, ρ_1 commutes with every deck transformation. It follows that $\delta\tau_p = \tau_{\delta(p)}$ for all p. Hence for $m \in M$ the geodesic segments τ_p, $p \in \pi^{-1}(m)$, all project to the same geodesic segment σ_m in M. Clearly σ_m is the unique shortest segment from m to A. Lifting to M_1 also shows that σ_m is the unique perpendicular from m to A, thus proving (1) and (2).

Define $h: M_1 \times I \to M_1$ by $h(p, t) = \tau_p(t)$. Since ρ_1 is continuous, so is h. But $\pi\tau_p = \sigma_{\pi(p)}$ implies $\pi h = H(\pi \times 1)$, and it follows that H is continuous, proving (3).

For $m, n \in M$ we can choose lifts $m_1, n_1 \in M_1$ such that $d(m, n) = d(m_1, n_1)$, by lifting a minimal geodesic from m to n. Then $d(m_1, n_1) \geq d(\rho_1(m_1), \rho_1(n_1)) \geq d(\rho(m), \rho(n))$, since both ρ_1 and π are distance-nonincreasing. This proves (4); the assertion for $K < 0$ follows from the corresponding assertion in the lemma.

We have defined a *closed geodesic* to be a nonconstant periodic geodesic; however, we use the same term for a nonconstant geodesic loop $\gamma: I \to M$ such that $\gamma'(0) = \gamma'(1)$.

3.5 COROLLARY. *Let S be a closed totally convex subset of M (complete, $K \leqq 0$). If $\gamma: I \to M$ is a closed geodesic, then γ is freely homotopic to a closed geodesic in A. If γ is not entirely in A, then all points of γ are at the same distance from A and the deformation takes place along a flat (finite length) cylinder with γ perpendicular to the meridians. Thus if $K < 0$, A contains every closed geodesic of M.*

Proof. Let $m = \gamma(0) = \gamma(1)$ and let τ_m be the perpendicular from m to A. Corresponding to the rectangle used in the proof of Lemma 3.2 (viewed as being in the simply connected covering of M) there is a rectangle in M whose longitudinal curves are the geodesic loops in the same free homotopy class as γ, based at the points of τ_m. Since the base curve, γ itself, is smooth the first variation $L^{2\prime}(0) = 2(\langle \tau'_m(0), \gamma'(1) \rangle - \langle \tau'_m(0), \gamma'(0) \rangle) = 0$. As before we have $L^{2\prime}(0) \leqq L^{2\prime}(v) \leqq L^{2\prime}(1) \leqq 0$, so L is constant, the second variation vanishes, and the deformation takes place along a flat cylinder. The longitudinal and transverse vector fields are again orthogonal, so the loops comprising the rectangle are smooth.

3.6 NOTE. The period of a closed geodesic does not have to be the same as that of its deformation into a totally convex subset. For example, in the flat Möbius strip the central circle is totally convex. The "parallel" closed geodesics have twice the period and can be deformed into the central one in the above manner.

3.7 COROLLARY. *If M (complete, $K < 0$) admits a convex function f without minimum, then there are no closed geodesics in M.*

Proof. Since f has no minimum, $\cap M^a$, for all values a of f, is empty. But by the preceding corollary any closed geodesic of M would have to be in this intersection.

3.8 REMARKS. (1) This result fails for $K \leqq 0$; for example, a (flat) circular cylinder admits such a function and has many closed geodesics.

(2) If M (complete, $K < 0$) admits a convex function having a minimum, then the minimum set contains all the closed geodesics of M.

If L is a level hypersurface of a function f on a manifold M, we denote by L^+ the *positive normal bundle* of L, that is, the set of all vectors x in the normal bundle $\perp L$ such that $xf \geqq 0$. Evidently L^+ is a manifold with boundary L.

3.9 PROPOSITION. *Let c be a noncritical value of a convex function f on M (complete, $K \leqq 0$). Let L^+ be the positive normal bundle of $L = f^{-1}(c)$. Then*

(1) *The exponential map of $\perp L$ is a diffeomorphism of L^+ onto ^cM leaving L fixed.*

(2) *The resulting geodesic deformation H of ^cM onto $L = f^{-1}(c)$ is length-nonincreasing. (If $K < 0$, H is length-decreasing except on curves in L.)*

Proof. If $x \in L^+$, then by definition $xf \geqq 0$. Since f is convex, f is nondecreasing on the geodesic ray with initial velocity x. Hence exp maps L^+ into ^cM. If $m \in {}^cM$, then since L is closed there is a shortest geodesic segment γ from L to m. Minimality

forces $\gamma'(0)$, the initial velocity of γ, to lie in L^+, hence $m \in \exp(L^+)$; that is, exp maps L^+ onto ^cM. That this mapping is one-one follows from the fact that perpendiculars to the totally convex set M^c are unique. Since geodesics do not realize distance to a submanifold beyond the first focal point, there can be no focal points along these perpendiculars. Hence exp is a diffeomorphism on L^+.

The normal bundle $\perp L$ is diffeomorphic to the product $L \times R^1$. The deformation H is given by the composition $\pi \exp^{-1}$, where $\pi: L^+ \to L$ is the product structure projection. If we give $\perp L$ the product metric, π is length-nonincreasing. Thus it suffices to prove that exp is length-nondecreasing. This is equivalent to showing that the lengths of L-Jacobi fields (those belonging to families of geodesics perpendicular to L) are nondecreasing along perpendiculars to M^c. Due to nonpositive curvature the length squared of any Jacobi field is a convex function of its geodesic's parameter. Thus it suffices to show that the first derivative, at the points of L, is nonnegative.

Let X be a nonzero vector field on L. Then X can be extended to ^cM so that it is a Jacobi field along each perpendicular; this extension is not unique, but depends also on the parametrization assigned to each perpendicular. However, every L-Jacobi field which is nonzero on L arises in this way. For such an extension there is a corresponding field of velocities V to the perpendiculars. On L we realize V as the product of the gradient G of f and a positive function g. Moreover, $[V, X] = 0$, since V and X may be realized along each perpendicular as the longitudinal and transverse fields of a rectangle. Thus $V\langle X, X\rangle = 2\langle \nabla_V X, X\rangle = 2\langle \nabla_X V, X\rangle = 2(Xg)\langle G, X\rangle + 2g\langle A_G X, X\rangle = 2g\nabla^2 f(X, X) \geq 0$.

If $K < 0$, the lengths of Jacobi fields are strictly convex, so exp increases the lengths of vectors which are not tangent to L and not annihilated by π.

3.10 REMARK. Part of what we have shown, $V\langle X, X\rangle \geq 0$, may be rewritten $\langle \nabla_X X, V\rangle \leq 0$, that is, L is *convex upward*.

3.11 COROLLARY. *If f is a nonconstant convex function on M (complete, $K \leq 0$), then any two level hypersurfaces of f are diffeomorphic.*

Proof. Let $b < c$ be noncritical values of f, and let $L = f^{-1}(b)$. By assertion (1) of the preceding proposition we can transfer the problem to L^+, setting

$$S = \exp^{-1}(f^{-1}(c)).$$

The projection π restricts to a one-one mapping of S onto L. If γ is a nonconstant geodesic ray initially normal to L, then f is strictly increasing on γ, hence γ' is never tangent to $f^{-1}(c)$. This means that a nonzero vertical tangent vector to L^+ cannot be tangent to S; hence $\pi: S \to L$ is a diffeomorphism.

3.12 THEOREM. *Let f be a convex function without minimum on a complete riemannian manifold with $K \leq 0$. Then M is diffeomorphic to $L \times R^1$, where $L = f^{-1}(b)$ is any level hypersurface of f.*

Proof. It suffices to exhibit a complete vector field V on M such that (a) each integral curve $\alpha: R \to M$ of V meets L for exactly one parameter value, and (b) V is never tangent to L. In fact, if ϕ is the flow of V, the restriction of ϕ to $L \times R^1$ is the required diffeomorphism. We construct V in two parts joined by a smoothing operation. On M^b, except in the neighborhood of L defined below, V coincides with the normalized gradient $U = \operatorname{grad} f / \|\operatorname{grad} f\|$. On ^bM, V will consist of velocities of unit speed geodesics normal to L.

Let N be a normal neighborhood of L in M; that is, N is the diffeomorphic image under exp: $\bot L \to M$ of a neighborhood N' of L in $\bot L$ such that if $x \in N'$, then $tx \in N'$ for all $0 \le t \le 1$. It follows from Proposition 3.9 that $N \cup {}^bM$ is also normal, being the diffeomorphic image of $N' \cup L^+$. Let W be the unit vector field on $N \cup {}^bM$ that is the image under exp of the obvious (positively directed) vertical vector field on $\bot L$. Geometrically, if $f(m) > b$, $W(m)$ is the velocity at m of the unique minimal unit-speed geodesic from L to m.

Now we smooth U and W to produce the required vector field V on M. Let g be the C^∞ function on N that measures directed distance to L, that is, g is zero on L and $\operatorname{grad} g = W$. There is a C^∞ function $r > 0$ on L such that for each $p \in L$ the closed ball of radius $r(p)$ lies in N. Extend r to N by constancy along integral curves of W. Then set $h = h^*(g/r)$, where $h^*: R \to [0, 1]$ is a C^∞ function that is 1 on $(-\infty, -1]$ and 0 on $[0, \infty)$.

On N, define $V = hU + (1-h)W$. We can suppose that $\langle U, W \rangle > 1/2$ on N. It follows that $Vg > 1/2$, for $Vg = \langle V, W \rangle = h\langle U, W \rangle + 1 - h > 1 - h/2$ and $h \le 1$.

Now we extend V from N to M by defining V to be U on $M^b - N$ and W on ^bM. Clearly V is a C^∞ vector field on the entire manifold M. Since the norm of V is at most one, V is complete. The integral curve of U starting at a point $m \in M^b$ eventually reaches N (by 2.3). Since $Vg > 1/2$ on N, the integral curve of V starting at m reaches L. The integral curve of V starting at $m \in {}^bM$ clearly reaches L at parameter value $t = -d(m, L)$. We know that $Uf > 0$ and $Wf > 0$, hence $Vf > 0$. Thus each integral curve of V meets L exactly once, and V is never tangent to L.

3.13 REMARK. Under the hypothesis of the theorem, M has at most two ends. If M has only one end, then the level hypersurfaces of any convex function without minimum are noncompact. Otherwise all are compact.

We shall say that a submanifold S of M (complete, $K \le 0$) *simplifies* M provided S is closed, totally convex, and every isometry of M carries S into itself. In view of 3.1 and 3.5, S then has the following properties:

3.14. S is a closed, connected, totally geodesic submanifold of M (hence is complete with nonpositive curvature).

3.15. exp: $\bot S \to M$ is a diffeomorphism, or equivalently, the inclusion $S \subset M$ induces an isomorphism of $\pi_1(S)$ onto $\pi_1(M)$.

3.16. Every closed geodesic of M is freely homotopic via a flat totally geodesic cylinder to a closed geodesic in S. (If $K < 0$, S contains every closed geodesic of M.)

3.17. For every $\phi \in I(M)$ the restriction $\phi | S$ is in $I(S)$.

3.18 REMARKS. (1) Evidently if S simplifies M, then to a considerable extent the topology and geometry of M are concentrated in M; e.g., they have the same homotopy type and (if $K<0$) the same closed geodesics.

(2) M may not have a simplifying submanifold $S \neq M$, e.g., if M is compact.

(3) In 3.16 the two closed geodesics need not have the same period; see 3.6.

(4) Simplification is transitive; that is, if S simplifies M (complete, $K \leqq 0$) and T simplifies S, then T simplifies M.

(5) Any nonempty intersection of simplifying submanifolds of M is again a simplifying submanifold (since any intersection of totally geodesic submanifolds is in particular a submanifold).

4. Construction of convex functions.   We consider some general methods of constructing convex and strictly convex functions.

On euclidean space ($K=0$) the matrix of second derivatives ($\partial^2 f/\partial x_i\,\partial x_j$) of a convex function is positive semidefinite and has "exact rows". Conversely, let (F_{ij}) be a symmetric, positive semidefinite matrix of functions such that $\sum F_{ij}\,dx_j$ is exact for $1 \leqq i \leqq n$; then there is a convex function f having $\partial^2 f/\partial x_i\,\partial x_j = F_{ij}$. One need only integrate twice and add an arbitrary linear function.

We now collect a number of examples constructed using the riemannian distance d.

4.1 THEOREM. *Let M be a complete simply connected riemannian manifold with $K \leqq 0$.*

(1) *If S is a closed, totally geodesic submanifold of M, the C^∞ function $f_S(m) = d^2(m, S)$ is convex. If $K<0$, then f_S is strictly convex on $M-S$; in fact, $\nabla^2 f_S(y, y)=0$ for $y \neq 0$ if and only if y is tangent to S.*

(2) *In (1), if S is a single point p, then f_p is strictly convex (for $K \leqq 0$).*

(3) *If S and T are closed totally geodesic submanifolds of M, then $f_S + f_T$ is convex with minimum set C the set of midpoints of common perpendiculars to S and T. (If S meets T, this means $C = S \cap T$.) Such perpendiculars join pairs $(s, t) \in S \times T$ that realize distance from S to T. If $K<0$, then $f_S + f_T$ is strictly convex if and only if S and T meet in at most one point.*

(4) *If μ is an arbitrary positive finite measure on M with compact support, then g, the mean square distance with respect to μ, is strictly convex: $g(p) = \int_M d^2(p, m)\mu$. The function g has a unique minimum point. (For example, μ could be the volume element of a compact submanifold of M.)*

Proof. The later cases are easy consequences of (1), which we prove first. For each $m \in M$ there is a unique perpendicular $\tau = \tau_m$ from m to S. Furthermore, $m \to -\tau'(1)$ is the inverse of exp: $\perp S \to M$, so f_S is C^∞. Moreover, $f_S(m) = |\tau_m|^2$.

Let $0 \neq y \in M_m$ and let β be the geodesic with initial velocity y. Then $(\nabla^2 f_S)(y, y)$ is the second variation of the squared arc length L^2 for the rectangle r given by $r(u, v) = \tau_{\beta(v)}(u)$. The integral term of $L^{2''}$ is the same as in Lemma 3.2, and is thus

nonnegative. The endpoint terms both vanish: if Y is the transverse vector field of r (so $Y = \partial r / \partial v$ and $Y(0) = y$), the endpoint terms are $\langle \nabla_Y Y, \tau' \rangle$ evaluated at $u = 0$ and $u = 1$. The term at $u = 0$ vanishes because the first transverse curve β is a geodesic; the term at $u = 1$ vanishes because the last transverse curve lies in S, which is totally geodesic and normal to $\tau'(1)$. Thus f_S is convex.

Now assume $\nabla^2 f_S(y, y) = 0$. Then we must have $\| \nabla_u Y \| = 0$, so Y is parallel along τ, and in particular is never zero.

If $K < 0$, then we must have $Y \wedge \tau' = 0$; hence $\tau'(1) = \alpha Y(1) \in S_{\tau(1)}$. But $\tau'(1)$ is normal to S, hence $\tau'(1) = 0$. It follows that $y \in S_m$ since τ is the constant curve $\tau = m \in S$.

In (2) it suffices to assume $K \leq 0$, since in this case $Y(1) \neq 0$ is a contradiction.

In (3) if S meets T, then $f_S + f_T$ has minimum value zero, so $C = S \cap T$, a closed totally geodesic submanifold. In general, let $m \in C$, and let σ and τ be the unique (nonconstant) perpendiculars from m to S and T, respectively. For the first variation of L^2, we compute

$$ yf_S = L^{2'}(0) = 2\langle Y, \sigma' \rangle]_0^1 = -2\langle y, \sigma'(0) \rangle. $$

Similarly for τ, we get $yf_T = -2\langle y, \tau'(0) \rangle$. Since $m \in C$ and C is the set of critical points of $f_S + f_T$, it follows that $\sigma'(0) + \tau'(0) = 0$. Thus σ and τ join to form a common perpendicular to S and T. Furthermore, since $|\sigma| = \| \sigma'(0) \|$ and $|\tau| = \| \tau'(0) \|$, m is the midpoint of this perpendicular. Reversing the argument, if m is such a midpoint, then $m \in C$.

If $\gamma : I \to M$ is a common perpendicular from S to T, then a first variation argument shows that $\gamma(1)$ is a critical point of $f_S | T$. But f_S is convex and T is totally geodesic, so $f_S | T$ is convex on T. Hence $\gamma(1)$ is a minimum point of $f_S | T$; that is, $d(\gamma(0), \gamma(1)) = d(S, T)$. The converse is immediate.

The $K < 0$ assertion in (3) follows from (2), so the proof of (3) is complete.

In (4), let γ be a geodesic with initial velocity $y \in M_m$. Then $(\nabla^2 g)(y, y) =$. $\int_M (\partial^2 / \partial t^2)(0)[d^2(\gamma(t), m)] \mu(m) = \int_M (f_m \circ \gamma)''(0) \mu(m) = \int_M (\nabla^2 f_m)(y, y) \mu(m)$. Thus by (2), g is strictly convex. Let q be in the support of μ and let r be the diameter of the support of μ. Then for any point p not in the compact ball of radius $2r$ at q, we have $g(p) > g(q)$. Thus g has at least one minimum point. But g is strictly convex, so as noted earlier, g can have at most one minimum point.

4.2 PROPOSITION. *If ϕ is an isometry of M, let f_ϕ be the square displacement function: $f_\phi(m) = d^2(m, \phi(m))$. Then f_ϕ is convex and exactly one of the following holds: (1) ϕ has a fixed point. Then C is the fixed point set of ϕ, a closed, connected, totally geodesic submanifold. (2) ϕ translates a geodesic. Then C consists of the images of all geodesics translated by ϕ (and is not necessarily a submanifold). (3) f_ϕ has no minimum.*

Furthermore, if $K < 0$, then f_ϕ is strictly convex except on the minimum set C, and ϕ translates at most one geodesic.

Proof. For $y \in M_m$, we compute yf_ϕ. Let α be the geodesic with initial velocity y; let r be the rectangle such that $r(\cdot, v)$ is the geodesic segment from $\alpha(v)$ to $\phi(\alpha(v))$; let Y be the associated vector field of r on $\gamma = r(\cdot, 0)$. Thus for $L(v) = |r(\cdot, v)|$ we have $L^2 = f_\phi \circ \alpha$. From the first variation formula, we compute $yf_\phi = L^{2\prime}(0) = 2\langle Y, \gamma'\rangle]_0^1 = 2(\langle \phi_*(y), \gamma'(1)\rangle - \langle y, \gamma'(0)\rangle)$. It follows that m is a *nonzero* critical point of f_ϕ if and only if ϕ translates the geodesic extension of γ, also denoted γ. Thus exactly one of the cases (1), (2), (3) above holds. Using the second variation formula we compute $(\nabla^2 f_\phi)(y, y) = \frac{1}{2} L^{2\prime}(0) = \int_0^1 \{\|Y'\|^2 - K(Y, \gamma')\|Y \wedge \gamma'\|^2\} \geq 0$. Hence f_ϕ is convex. In case (1) it is well known (for any M) that the fixed point set of an isometry is a closed totally geodesic submanifold. Here C is connected, since C is totally convex.

Now suppose $(\nabla^2 f_\phi)(y, y) = 0$ for $y \neq 0$. The formula above implies that Y is parallel on γ. Furthermore, assuming from now on that $K < 0$, it implies $\|Y \wedge \gamma'\| = 0$. If $\gamma' = 0$, then m is a fixed point of ϕ. If m is not a fixed point, then $Y = \mu \gamma'$ for some number $\mu \neq 0$. But $Y(0) = y$ and $Y(1) = \phi_*(y)$, hence in particular, $\phi_*(\gamma'(0)) = Y(1)/\mu = \gamma'(1)$; so ϕ translates γ. Also $Y = \mu \gamma'$ shows that γ' spans the nullspace of $\nabla^2 f_\phi$ at m—and in fact at every point of $\gamma(R)$. Thus $\gamma(R)$ has a neighborhood on which $\nabla^2 f_\phi$ is nondegenerate except on $\gamma(R)$. This means that $\gamma(R)$ is a component of C, hence by connectedness equals C.

4.3 REMARKS. (1) Some of the results above extend easily to the case where M is not simply connected. For example, if S is a closed totally convex submanifold, then the function $f_S(m) = d^2(m, S)$ is C^∞ since it coincides with $\|\exp^{-1}(m)\|^2$. Then (1) and (3) of Theorem 4.1 hold with totally geodesic strengthened to totally convex. In Proposition 4.2 simple connectedness was used in giving the geodesic from m to $\phi(m)$ uniquely and smoothly. If $\phi = \phi_1$ is homotopic to the identity via a continuous homotopy $\{\phi_t\}$, then the geodesic in the homotopy class of the curve $t \to \phi_t(m)$ is another unique and smooth function of m. Thus the proposition follows as before, except that "fixed point" is no longer the appropriate notion in (1). For example, the identity on a hyperboloid of one sheet may be viewed as one complete revolution, in which case the fixed point set is the whole surface, but the minimum point set C consists of the points on the closed geodesic. The notion used in (1) should be "points of zero displacement".

(2) In Theorem 4.1 (4) we have given the nontrivial part of the proof of a well-known theorem of Cartan: *If G is a compact group of isometries of a complete simply connected manifold M with $K \leq 0$, then the members of G have a common fixed point.* In fact, let μ be the volume element of an orbit under G. Then the function g of the theorem is invariant under G, so every member of G fixes the unique minimum point of g.

(3) The above results can be strengthened by replacing square distances by distances, although in doing so we must deal with continuous rather than C^∞ functions. A continuous function $f: M \to R$ is *convex* if for every geodesic segment $\gamma: I \to M$ from any m to any n and for every $v \in I$, we have $f(\gamma(v)) \leq (1 - v)f(m) + vf(n)$.

This is consistent with the definition for C^∞ functions. We note that if f is also nonnegative, then we may square the inequality and use the fact that $(1-v)^2 \leq 1-v$, etc., to obtain $f^2(\gamma(v)) \leq (1-v)f^2(m) + vf^2(n)$. Thus if f is nonnegative and convex so is f^2. Hence Proposition 4.7 below is a generalization of Theorem 4.1 (1). Before stating and proving it we need some facts about continuous convex functions, some of which are well known and will be given without proof.

4.4 LEMMA. *Convexity is a local property; that is, if for every $m \in M$, f is convex in some neighborhood of m, then f is convex.*

4.5 LEMMA. *Let $f: (-\alpha, \alpha) \to R$ be continuous (here α may be $+\infty$). (1) If f is convex, then the right and left derivatives f'_+ and f'_- exist on all of $(-\alpha, \alpha)$ and are nondecreasing. Hence the second derivative exists almost everywhere and is nonnegative.*

(2) *If f is convex on $(-\alpha, 0]$ and $[0, \alpha)$, and $f'_-(0) \leq f'_+(0)$, then f is convex.*

(3) *If 0 is a minimum point of f and f is convex on $(-\alpha, 0]$ and $[0, \alpha)$, then f is convex.*

4.6 LEMMA. *Let M be simply connected, complete, with $K \leq 0$. (1) If S is a closed totally geodesic submanifold of M, then distance to S is a convex function on M. (2) If S is a geodesic segment in M, then distance to S is a convex function on M.*

Proof. (1) Let r be the distance to S. Then by Theorem 4.1 (1), r^2 is C^∞ and convex. Moreover, r is C^∞ on $M-S$ and there we have $\nabla^2 r^2 = 2\nabla(r\,dr) = 2(dr)^2 + 2r\nabla^2 r$. Thus $\nabla^2 r$ is positive semidefinite on the hyperplanes annihilated by dr. But the direction $E_1 = \operatorname{grad} r$ is a null direction for $\nabla^2 r$, because $E_1 r = 1$ is constant and $\langle \nabla_{E_2} E_1, E_1 \rangle = 0$ for any vector field E_2. Thus $\nabla^2 r(E_2, E_1) = E_2 1 - (\nabla_{E_2} E_1) r = 0$. This shows that $\nabla^2 r$ is positive semidefinite on $M-S$. But r takes its minimum value on S, so r is convex on any geodesic meeting S by Lemma 4.5 (3). Hence r is convex on M.

(2) We divide M into three parts, depending on whether the nearest point on S is an interior point or one or the other endpoint. In the interior of any of these parts the distance r to S is convex by (1). At the endpoints of S, r takes its minimum, and so is convex on any geodesic through one of the endpoints by Lemma 4.5 (3). In a neighborhood of another boundary point of the three regions, r is C^1 by the formula for first variation of arc length. Thus on a geodesic which is transverse to the boundary r is convex by Lemma 4.5 (2). A geodesic which is tangent to the boundary is the limit of transverse geodesics. Thus r is convex on all geodesics.

4.7 PROPOSITION. *Let M be a complete riemannian manifold with $K \leq 0$ and S a closed totally convex subset of M. Then the distance to S is a continuous convex function on M.*

Proof. Let $\pi: M_1 \to M$ be the simply connected covering space of M, $S_1 = \pi^{-1}(S)$, r the distance to S, and r_1 the distance to S_1. Then by Proposition 3.4 the segments

whose lengths realize r are unique, so their lifts to M_1 are the segments whose lengths realize r_1. It follows that $r \circ \pi = r_1$ and we need only prove that r_1 is convex. Thus we assume henceforth that M is simply connected.

Let $\gamma: I \to M$ be a geodesic segment from m to n. The nearest points in S to m and n are unique, say $\rho(m)$ and $\rho(n)$. Let $\tau: I \to M$ be the geodesic segment from $\rho(m)$ to $\rho(n)$. Since S is totally convex $\tau(I) \subset S$. Let $f(v)$ be the distance from $\gamma(v)$ to $\tau(I)$, so f is convex by Lemma 4.6 (2). Thus we have

$$r(\gamma(v)) \leqq f(v) \leqq (1-v)f(0) + vf(1) = (1-v)r(\gamma(0)) + vr(\gamma(1)),$$

that is, r is convex.

We recall some facts about the simply connected covering $\pi: M_1 \to M$ of a manifold M. The fundamental group $\pi_1(M)$—without base point—can be considered to be the deck transformation group D of the covering. If $\delta \in D$, let $F(\delta)$ be the set of all loops in M of the form $\pi \circ \alpha$, where $\alpha: I \to M_1$ is a curve such that $\delta(\alpha(0)) = \alpha(1)$. Then $F(\delta)$ is an entire free homotopy class of loops in M. If δ is in the *center* of D, then for each $m \in M$, the loops of $F(\delta)$ at m constitute a single element of $\pi_1(M, m)$.

4.8 LEMMA. *Let M be complete, with $K \leqq 0$. For each element δ in the center of $\pi_1(M)$ there is a convex function f on M whose minimum set consists of the images of all closed geodesics in the free homotopy class $F(\delta)$. (Explicitly, $f(m) = |\gamma_m|^2$, where γ_m is the unique geodesic loop of $F(\delta)$ at $m \in M$.) If $F(\delta)$ contains no closed geodesics, then f is strictly convex.*

Proof. Let f_1 be the square displacement function of δ, as in Proposition 4.2. By hypothesis δ commutes with every element of D, hence f_1 can be factored through $M: f_1 = f \circ \pi$. Thus f is convex, and clearly f is the function defined in the statement of this lemma. Since δ has no fixed points, Proposition 4.2 says that the minimum set C_1 of f_1 consists of the images of all geodesics translated by δ. It follows, using the remarks above, that the minimum set $C = \pi(C_1)$ of f consists of the images of all closed geodesics in $F(\delta)$. If there exist no closed geodesics in $F(\delta)$, then $C_1 = \pi^{-1}(C)$ is empty, so f_1 is strictly convex and so is f.

For negative curvature we get the following stronger result:

4.9 THEOREM. *Let M be a complete riemannian manifold with $K < 0$. If $\pi_1(M)$ has nontrivial center, then*

(1) *M is diffeomorphic to a product $L \times R^1$ unless M is a Möbius band.*

(2) *If there is a closed geodesic γ in M, then γ is simply closed and $S^1 = \gamma(R)$ simplifies M. In particular, M is a vector bundle over S^1, and γ is (up to parametrization) the unique closed geodesic in M.*

Proof. Let $\delta \neq 1$ be a central element of $\pi_1(M)$. In the context of the preceding proof, Proposition 4.2 implies that δ translates at most one geodesic of M_1, since $K < 0$.

If δ does not translate a geodesic, then f_1, hence f, is strictly convex without minimum. Thus by Theorem 3.12, M is diffeomorphic to $L \times R^1$, and by Corollary 3.7, M contains no closed geodesics.

If δ translates a (necessarily unique) geodesic γ, then the minimum set C of the function f is the image of the closed geodesic $\pi \circ \gamma$ in M. Since C is totally convex, it follows that $\pi \circ \gamma$ is simply closed; that is, C is a circle. By Corollary 3.5, $\pi \circ \gamma$ is essentially the only closed geodesic in M. Thus C is invariant under every isometry of M, so that C simplifies M. Simplification implies that exp: $\perp C \to M$ is a diffeomorphism; this structures M as a vector bundle over the circle C. If M is orientable as a manifold, or equivalently as a vector bundle, then M is diffeomorphic to $C \times R^{d-1}$. Otherwise, M is diffeomorphic to $B \times R^{d-2}$, where B is a Möbius band containing C. Thus assertion (1) holds also in this case.

Note that the Möbius band B is a genuine exception to assertion (1) above, since B admits a complete riemannian structure having (constant) negative curvature.

5. **Mobility.**   Let M be a riemannian manifold. We say that a point m of M is *mobile* provided there exists a Killing vector field X on M such that $X(m) \neq 0$. Otherwise m is *immobile*. Let $P(M)$ be the set of all immobile points of M. Then M is *mobile* provided $P(M)$ is empty (every point mobile), and M is *immobile* provided $P(M) = M$ (every point immobile). Of course, M can be neither mobile nor immobile.

If M is complete, then the Lie algebra $i(M)$ of all Killing vector fields on M is the Lie algebra of the connected component $I_0(M)$ of the isometry group $I(M)$ of M. It follows that the set $P(M)$ of immobile points of M is exactly the set of common fixed points of all elements of $I_0(M)$. In particular, the following are equivalent: M immobile, $i(M) = 0$, $I(M)$ discrete.

We mention a few examples. A zero-dimensional manifold is trivially immobile; a one-dimensional manifold is mobile. A homogeneous riemannian manifold M is mobile, since $I_0(M)$ is transitive on M. A compact manifold M with $K < 0$ is immobile, since it is known that $I(M)$ is finite. We can prove this as follows: For every Killing field X on M, the function $\|X\|$ must have a maximum. Since it is convex (see 5.5 below), it must be constant. But then every integral curve of X is a geodesic, contradicting 5.5 unless $X = 0$. Thus $I(M)$ is discrete and compact, hence finite. (Compare [3].)

If ϕ is an isometry of a complete riemannian manifold M, the fixed point set $F(\phi)$, if nonempty, is a closed totally geodesic submanifold of M [5]. Also a nonempty intersection of closed, totally geodesic submanifolds is again such a submanifold. Thus if $P(M)$ is nonempty (that is, M is not mobile), then $P(M)$ is a closed totally geodesic submanifold, since $P(M)$ is the intersection of the fixed point sets $F(\phi)$ for all $\phi \in I_0(M)$. In general, $P(M)$ need not be connected. We assert that no component of $P(M)$ is a hypersurface. In fact if $P(M)$ contains a

connected hypersurface S, then each $\phi \in I_0(M)$ leaves S pointwise fixed; thus, for $m \in S$, $(S_m)^\perp$ is one-dimensional and invariant under the rotation ϕ_{*m}. Hence ϕ_{*m} is the identity map of M_m and ϕ is the identity map of M. But then $P(M) = M$.

We get a considerably stronger description of $P(M)$ when M has nonpositive curvature.

5.1 PROPOSITION. *Let M be a complete riemannian manifold with $K \leq 0$. If $P = P(M)$ is nonempty, then P is a submanifold that simplifies M (and is not a hypersurface). Thus P has properties 3.14–17; in particular, it is connected.*

Proof. We must show that P is totally convex and invariant under all isometries of M. Let $\tau: I \to M$ be a geodesic segment joining points $p, q \in P$. If $\phi \in I_0(M)$ there is a homotopy $\{\phi_t\}$ in $I_0(M)$ from the identity map ϕ_0 to $\phi_1 = \phi$. Since each ϕ_t fixes p and q, it follows that $\phi \circ \tau$ is fixed-endpoint homotopic to τ. Then $K \leq 0$ implies $\phi \circ \tau = \tau$; that is, every $\phi \in I_0(M)$ leaves τ pointwise fixed. Hence τ lies in P.

To prove the isometry condition, let $\mu \in I(M)$. If $\phi \in I_0(M)$, then $\mu^{-1}\phi\mu \in I_0(M)$. Thus if $p \in P$, we have $\mu^{-1}\phi\mu(p) = p$. Hence $\mu(p)$ is a fixed point of every $\phi \in I_0(M)$, so $\mu(p) \in P$.

We now show how to reduce any M (complete, $K \leq 0$) to the mobile or immobile case, up to simplification. If $P(M)$ is not empty, then $P^2(M) = P(P(M))$ is well defined. When M is complete, $K \leq 0$, so are the nonempty iterates $P^i(M)$. Thus

5.2 COROLLARY. *If M is a complete riemannian manifold with $K \leq 0$, there is a smallest integer k such that $P^k(M)$ is either mobile or immobile. Furthermore, $P^k(M)$ simplifies M, and $k \leq \frac{1}{2} \dim M$.*

Proof. The simplification assertion follows from the preceding proposition, since simplification is transitive.

To prove the last assertion, note that for $1 \leq i \leq k$, $P^{i-1}(M) \neq P^i(M)$. We have seen that $P^i(M)$ cannot be a hypersurface in $P^{i-1}(M)$. Hence by recursion $0 \leq \dim P^i(M) \leq \dim M - 2i$, and the result follows.

We call this integer k, the smallest such that $P^k(M)$ is either mobile or immobile, the *mobility grade* of M. When $P^k(M)$ is mobile, that is, $P^{k+1}(M)$ is empty, we say that M is *ultimately mobile*. When $P^k(M)$ is immobile, that is, equals $P^{k+1}(M)$, we say that M is *ultimately immobile*.

For example, if $P^k(M)$ is compact (in particular, zero-dimensional), then M is ultimately immobile. If $\dim P^k(M) = 1$, then M is ultimately mobile, and in fact is diffeomorphic to one of three types (for each dimension ≥ 3): simply connected (R^d), the oriented vector bundle over S^1 ($S^1 \times R^{d-1}$), and the nonoriented vector bundle over S^1 ($B \times R^{d-2}$, $B = $ Möbius band). If $\dim P^k(M) \geq 2$, no conclusions can be drawn (see Corollary 8.4).

In §8 we show that there are complete $K < 0$ manifolds M with arbitrary mobility grade $k \leq \frac{1}{2} \dim M$.

We recall some formulas for Killing vector fields.

5.3 LEMMA. *Let* $f = \|X\|$, *where* X *is a Killing vector field on a riemannian manifold* M. *If* $v \in M_m$ *and* $f(m) \neq 0$, *then*

(1) $vf = \langle \nabla_v X, X(m) \rangle / f(m) = -\langle (\nabla_X X)(m), v \rangle / f(m)$; *hence* $\operatorname{grad} f^2 = -2\nabla_X X$.

(2) $f^3(m) \nabla^2 f(v, v) = \|\nabla_v X \wedge X(m)\|^2 - f^2(m) K(v, X(m)) \|v \wedge X(m)\|^2$.

(3) $\nabla^2 f^2(v, v) = 2\{\|\nabla_v X\|^2 - K(v, X(m)) \|v \wedge X(m)\|^2\}$. *Here we need not assume* $f(m) \neq 0$.

Proof. (1) Since X is a Killing field, A_X is skew-symmetric at each point. Thus

$$vf = v\langle X, X \rangle^{1/2} = \langle \nabla_v X, X(m) \rangle / f(m) = \langle A_X v, X(m) \rangle / f(m)$$
$$= -\langle A_X X(m), v \rangle / f(m) = -\langle \nabla_X X(m), v \rangle / f(m).$$

We also observe that $\operatorname{grad} f^2 = 2f \operatorname{grad} f$.

(2) Let γ be the geodesic with initial velocity v, so $\nabla^2 f(v, v) = (f \circ \gamma)''(0)$. Since the restriction of X to γ is a Jacobi field on γ, the result follows from the Jacobi differential equation and a straightforward computation. The proof of (3) is by the same means as (2), only simpler.

The following consequence is also well known.

5.4 COROLLARY. *A point* $p \in M$ *is a* [*nonzero*] *critical point of* $\|X\|^2$ *if and only if the integral curve* γ *of* X *such that* $\gamma(0) = p$ *is a* [*nonconstant*] *geodesic.*

We now prove an infinitesimal analogue of Proposition 4.2.

5.5 PROPOSITION. *Let* X *be a Killing vector field on* M (*complete,* $K \leq 0$). *Then the function* $f = \|X\|$ *is continuous convex, and there are exactly three mutually exclusive possibilities*:

(1) X *vanishes at some point of* M. *The set* $C = M^0$ *of zeros of* X *is a closed, totally convex submanifold of even codimension.*

(2) X *has a nonconstant geodesic integral curve* γ. *The union* C *of the images of all such geodesics is a closed, totally convex set.*

(3) f *has no minimum points.*

Furthermore: if $K < 0$, *then* f *is strictly convex on* $M - C$. *In* (2) γ *is unique except for parametrization,* $C = \gamma(R)$ *is a closed totally convex submanifold, and* $\nabla^2 f$ *is degenerate only on vectors tangent to* C.

Proof. It is clear from Lemma 5.3 that f is convex wherever it is nonzero. If γ is a geodesic on which f vanishes at more than one point, then f vanishes identically on γ, so $f \circ \gamma$ is certainly convex. If f vanishes at only one point of a geodesic γ, then $f \circ \gamma$ is convex by Lemma 4.5 (3). Thus f is convex everywhere. Moreover, we see from Corollary 5.4 that the initial statements in (1), (2), (3) are mutually exclusive and all inclusive.

Let C be the minimum set of f, so C is closed and totally convex (in particular, connected). In (1), $C = M^0$, and M^0 is a submanifold, since it is the intersection of the fixed point sets of the isometries $\exp tX$ for all t. If $m \in M^0$, then near m, M^0 coincides with the fixed point set of a single isometry $\phi = \exp tX$, t small. Since the

differential map ϕ_{*m} is a rotation of M_m, its $(+1)$-eigenspace V has even codimension in M_m. But V is the tangent space of M^0 at m; hence M^0 has even codimension in M.

In (2), the character of C follows from Corollary 5.4.

Now suppose $K<0$. If $0 \neq v \in M_m$, $f(m) \neq 0$, and $\nabla^2 f(v, v) = 0$, then by Lemma 5.3 (2), $v = \alpha X(m)$ for some nonzero number α, and $\nabla_v X \wedge X(m) = 0$. But $\nabla_X X$ is orthogonal to X since A_X is skew-symmetric, so $\nabla_X X(m) = 0$ and we have case (2), with $m \in C$. Since the nullspace of $\nabla^2 f$ is 1-dimensional, f has no other minimum points in a tubular neighborhood of the (geodesic) integral curve γ of X. Since C is connected, we conclude that $C = \gamma(R)$.

All three types of Killing fields can be exhibited on hyperbolic space. In case (1) we call X an *infinitesimal rotation about* M^0 (which in H^3 must be a geodesic). In case (2), X is an *infinitesimal spiral translation along* γ, or just an *infinitesimal translation* if $A_X \circ \gamma = 0$. In case (3) we can consider X to be an *infinitesimal rotation about infinity*.

Note that Propositions 2.2 and 5.5 imply that if M (complete, $K<0$) has finite volume, it is immobile.

6. Ultimately mobile manifolds.   By a *Killing geodesic* we mean a nonconstant geodesic that is an integral curve of a Killing field. Note that *if M (complete, $K \leq 0$) contains a Killing geodesic, then M is mobile*, since by Proposition 5.5 there is a nonvanishing Killing field on M. As a partial converse we have

6.1 LEMMA. *If M (complete, $K<0$) is mobile, then every closed geodesic in M is a Killing geodesic.*

Proof. Let γ be a closed geodesic in M. Since $P(M)$ is empty, there exists a Killing field X on M whose restriction $X \circ \gamma$ to γ is not identically zero. If $X \circ \gamma$ is not tangent to γ, then for small t, $\exp tX$ carries γ to a distinct closed geodesic that is freely homotopic to γ. But this contradicts $K<0$, hence $X \circ \gamma(s) = h(s)\gamma'(s)$. Because $\nabla_X X$ is orthogonal to X, h is constant (nonzero), so γ is an integral curve of X/h.

(Further conclusions in this direction are included in Theorem 6.4, since, of course, mobile implies ultimately mobile.)

6.2 LEMMA. *Let γ be a Killing geodesic in M (complete, $K<0$).*
(1) *If γ is one-one, then M is simply connected.*
(2) *If γ is closed, then the circle $\gamma(R)$ simplifies M.*

Proof. Let X be a Killing field of which γ is a nonconstant integral curve. By Proposition 5.5, $P = \gamma(R)$ is the minimum set of the convex function $\|X\|$, so P is a closed totally convex submanifold. Thus by Lemma 3.1, $\exp: \perp P \to M$ is a diffeomorphism. A vector bundle over R^1 is trivial, which proves (1). In case (2) it follows from Corollary 3.5 that γ is essentially the only closed geodesic in M, so every isometry of M leaves γ invariant. Thus the circle $\gamma(R)$ simplifies M.

We have seen that if M admits a strictly convex function, then M contains no closed geodesics. The converse holds in the following special case:

6.3 Lemma. *If M (complete, $K<0$) is mobile, but contains no closed geodesics, then M admits a strictly convex function.*

Proof. If M contains a *nonclosed* Killing geodesic, then M is simply connected (6.2 (1)) and admits a strictly convex function (4.1 (2)). Thus we can assume that M contains no Killing geodesics.

Let $X_1, \ldots, X_n$ be a basis for the space $i(M)$ of Killing fields on M. By Proposition 5.5, the function $f = \sum \|X_i\|^2$ is strictly convex on $M - \bigcap C_i$, where C_i is the minimum point set of $\|X_i\|$. However, none of the X_i fall in case (2) of Proposition 5.5, so C_i consists of the points at which X_i vanishes. Thus $\bigcap C_i = P(M)$, which is empty, and f is strictly convex on M.

The following theorem has the same conclusion as Theorem 4.9.

6.4 Theorem. *Let M be a complete riemannian manifold with $K<0$. If M is ultimately mobile, then*

(1) *M is diffeomorphic to a product $L \times R^1$ unless M is a Möbius band.*

(2) *If there is a closed geodesic γ in M, then γ is simply closed and $S^1 = \gamma(R)$ simplifies M. In particular, M is a vector bundle over S^1, and γ is (up to parametrization) the unique closed geodesic in M.*

Proof. Let $Q = P^k(M)$, where k is the mobility grade of M. By hypothesis Q is mobile and by Proposition 5.1, Q is a closed totally convex submanifold of M.

Suppose first that there is a closed geodesic γ in M. By Corollary 3.5, γ lies in Q, and by Lemma 6.1, γ is a Killing geodesic of Q. Hence by Lemma 6.2, $S^1 = \gamma(R)$ simplifies Q. By Corollary 5.2, Q simplifies M, hence by transitivity, S^1 simplifies M. As mentioned earlier, the uniqueness of γ then follows from Corollary 3.5, so (2) is proved. Since S^1 simplifies M, it follows as in the proof of Theorem 4.9 that M is diffeomorphic to either $S^1 \times R^{d-1}$ or $B \times R^{d-2}$ (B a Möbius band). Thus (1) holds in this case.

Now suppose that M does not contain a closed geodesic. Thus Q does not contain a closed geodesic (since Q is totally geodesic). By Lemma 6.3, Q admits a strictly convex function f. If f has a minimum, then Q is simply connected. Hence M is simply connected and (1) holds trivially in this case. Finally, if f has no minimum, then by 3.12, Q is diffeomorphic to a product $L \times R^1$. Since Q simplifies M, M is the total space of a vector bundle β over $L \times R^1$. It follows that M is diffeomorphic to $N \times R^1$, where N is the total space of $\beta | L$.

6.5 Corollary (to 4.9 and 6.4). *Each of the following conditions implies that M (complete, $K<0$) is ultimately immobile and $\pi_1(M)$ has trivial center:*

(1) *M contains two distinct closed geodesics.*

(2) *M contains a closed but not simply closed geodesic (for example, a figure eight).*

(3) M is not diffeomorphic to a product $N \times R^1$ or to a Möbius band.

(4) M has at least three ends.

We now derive some relations between mobility and the isometry group. For this we will use some facts about the isometry group which are either well known or easily derived (cf. [6], pp. 47–48). If S is a closed riemannian submanifold of M (complete), then the set of isometries $E(S, M)$ which can be obtained as restrictions of isometries of M forms a closed subgroup of $I(S)$. Moreover, if S is invariant under $I(M)$, then the restriction homomorphism $I(M) \to E(S, M) \subset I(S)$ is a C^∞ map and its kernel K is isomorphic to a closed subgroup of $O(d-n)$, where $d = \dim M$, $n = \dim S$. (K is isomorphic to the subgroup of the linear isotropy group at a point of S which leaves the tangent space of S fixed.) Thus we have

6.6 PROPOSITION. (1) *If S is an n-dimensional submanifold of M invariant under $I(M)$, then*

$$\dim I(M) \leqq \dim I(S) + \dim O(d-n).$$

(2) *$I(M)$ is compact if and only if some (or every) orbit under $I(M)$ is compact.*

(3) *The subgroup of $I(M)$ leaving a compact submanifold invariant is compact.*

(4) *$I(M)$ is compact if there is a compact invariant submanifold.*

(5) *The same statements are true with $I(M)$ replaced by $I_0(M)$.*

(6) *If an isometry $\phi \in I(M)$ leaves a compact submanifold S invariant, then either ϕ has finite order or there is a nontrivial Killing field on M. (The subgroup $\{\phi^k\}$ lies in a compact subgroup of $I(M)$, so if $\{\phi^k\}$ is not finite, its closure contains a one-parameter subgroup.)*

The hypothesis of (1) above is satisfied when S simplifies M. Accordingly, when we may conclude by Theorem 4.9 or 6.4 that M is simplified by a circle we obtain further:

6.7 COROLLARY. *Suppose that M (complete, $K<0$) is ultimately mobile or $\pi_1(M)$ has nontrivial center. If M contains a closed geodesic, then $I(M)$ is compact and has dimension at most $(d^2 - 3d + 4)/2$.*

There is no corresponding result in the ultimately mobile case when M does not contain a closed geodesic, since this class of manifolds includes the hyperbolic spaces. However, we do have a sort of reverse result:

6.8 COROLLARY. *If M (complete, $K<0$) has $I_0(M)$ noncompact, then M is ultimately mobile and contains no closed geodesics.*

This is more evident in its contrapositive form:

6.9 COROLLARY. *If M (complete, $K<0$) is ultimately immobile, then $I_0(M)$ is compact.*

Proof. By hypothesis M is simplified by $Q = P^k(M)$ with $I(Q)$ discrete. Thus $\rho: I(M) \to I(Q)$ has a compact kernel containing $I_0(M)$ as a closed subgroup.

7. **Warped products.** This notion (and a generalization given at the end of this section) will be used to construct a large class of complete manifolds of negative curvature. A riemannian product cannot, of course, have negative curvature, but we can achieve it (Theorem 7.5) by altering the product riemannian structure as follows: Let B and F be riemannian manifolds and $f > 0$ a differentiable function on B. Consider the product (differentiable) manifold $B \times F$ with its projections $\pi \colon B \times F \to B$ and $\eta \colon B \times F \to F$. The *warped product* $M = B \times_f F$ is the manifold $B \times F$ furnished with the riemannian structure such that

$$\|x\|^2 = \|\pi_*(x)\|^2 + f^2(\pi m)\|\eta_*(x)\|^2$$

for every tangent vector $x \in M_m$.

For example, every surface of revolution (not crossing the axis of revolution) is isometric to a warped product, with B the generating curve, F the circle of unit radius, and $f(b)$ the distance from $b \in B$ to the axis of revolution.

7.1 REMARKS. Let $M = B \times_f F$.

(1) For each $p \in F$ the restriction of π to the *horizontal leaf* $\eta^{-1}(p)$ is an isometry onto B.

(2) For each $b \in B$ the restriction of η to the *vertical fiber* $\pi^{-1}(b)$ is a homometry onto F with scale factor $1/f(b)$.

(3) Since π_* is obviously length-nonincreasing, π is length-nonincreasing on curves. It follows from the local minimizing character of geodesics that a geodesic (intrinsic) of a horizontal leaf is a geodesic of M; that is, the horizontal leaves are totally geodesic.

(4) If ϕ is an isometry of F, then $1 \times \phi$ is an isometry of M. Thus if F is mobile (e.g., 1-dimensional), then M is mobile.

(5) If ψ is an isometry of B such that $f = f \circ \psi$, then $\psi \times 1$ is an isometry of M. Thus if X is a nonvanishing Killing field on B such that $Xf = 0$, then M is mobile. In particular, $B \times_k F$ is mobile, where $k = \|X\|^2$.

7.2 LEMMA. *$M = B \times_f F$ is complete if and only if B and F are complete.*

Proof. If M is complete, then a Cauchy sequence in B or F imbeds in a (horizontal) leaf or a (vertical) fiber as a Cauchy sequence, and hence converges.

If B and F are complete, let $\{m_i\}$ be a Cauchy sequence in M, with $m_i = (b_i, p_i)$. Let α_{ij} be a curve from m_i to m_j in M having length at most $2d(m_i, m_j)$. We can assume that all projections $\pi \circ \alpha_{ij}$ lie in a compact region in B, and on this we have $f \geq c > 0$. Thus the speed of α_{ij} at each point is at least c times the speed of $\pi \circ \alpha_{ij}$. Thus $d(p_i, p_j) \leq (2/c)d(m_i, m_j)$, showing that $\{p_i\}$ is Cauchy and hence convergent.

Since π is distance-nonincreasing, $\{b_i\}$ is also Cauchy, hence convergent. Thus $\{m_i\}$ is convergent, and M is complete.

We now turn to the study of the relations between the curvature (tensor and sectional) of M and that of B and F. The decomposition of vectors into horizontal and vertical parts, the second fundamental forms of the fibers, and the warping

function f and its gradient G all play a role in this development. The horizontal and vertical parts of a vector field A on $M = B \times_f F$ will be denoted by $\mathcal{H}A$ and $\mathcal{V}A$. A vector field X on B will be identified with the horizontal vector field on M that is π-related to X; a vector field V on F will be identified with the vertical η-related field. The function f will be identified with $f \circ \pi$. This introduces two meanings for the gradient G of f, but these coincide under a previous identification. In fact, $f \circ \pi$ is constant on fibers, so grad $(f \circ \pi)$ is horizontal; also for X on B (and M) we have $\langle X, \text{grad } (f \circ \pi) \rangle = X(f \circ \pi) = Xf \circ \pi = \langle X, G \rangle$. Finally, we shall sometimes identify a function g on F with $g \circ \eta$ on M.

The fact that the leaves of M are totally geodesic and naturally isometric to B lets us consider the structural objects on M as extensions of those on B. Thus we will make no notational distinction between the riemannian metric, curvature tensor or sectional curvature of B and M. For the fibers and F it is a different matter, so we make the following conventions. The riemannian metric on F will be denoted by $(\ ,\ )$, and for vector fields V, W on F we denote their inner product on M by $\langle V, W \rangle = f^2(V, W)$. The curvature tensor of F is denoted by S; that of M by R. The sectional curvature of F is denoted by L; that of M by K. The covariant differential operator on F is D; that of M is ∇.

We express the second fundamental form data of the fibers by the (1, 2) tensor T [7], which measures the difference between ∇ and the covariant differential of the fibers. Of course, such differences are defined only with respect to tangent vectors to the fibers, but we extend to arbitrary vectors by first projecting them onto the fibers. Thus for vector fields A, B on M we have

$$T_A B = \mathcal{H} \nabla_{\mathcal{V}A}(\mathcal{V}B) + \mathcal{V} \nabla_{\mathcal{V}A}(\mathcal{H}B).$$

At each point $m \in M$, T_A is a skew-symmetric linear operator on M_m reversing the horizontal and vertical subspaces. Moreover the symmetry of the second fundamental form is expressed by the fact that for vector fields V, W on F, $T_V W = T_W V$.

7.3 LEMMA. *Let X, Y be vector fields on B, and V, W vector fields on F. Then*
(1) *$\nabla_X Y$ is the same on either B or M.*
(2) *$\nabla_X V = \nabla_V X = T_V X = (Xf/f)V$.*
(3) *$\mathcal{H}(\nabla_V W) = T_V W = -f(V, W)G = -(\langle V, W \rangle/f)G$.*
(4) *$\mathcal{V}(\nabla_V W) = D_V W$.*

Note that (a) the last equality in (2) and the equalities in (3) are *tensorial*, hence are valid for arbitrary horizontal X and vertical V, W; (b) by (3) each (vertical) fiber is totally umbilic, with normal curvature vector $-(1/f)G$. In particular, the fibers at critical points of f (that is, where $G = 0$) are totally geodesic and other fibers are not.

Proof. (1) is equivalent to the fact that leaves are totally geodesic and isometric to B. For (2), since $[X, V] = 0$ we have $\nabla_X V = \nabla_V X$. Since $\langle X, Y \rangle$ is constant on

fibers, $V\langle X, Y\rangle = 0$, and the remaining terms in the Koszul formula [1, p. 166] for $\langle \nabla_V X, Y\rangle$ are zero for similar reasons. Thus $\nabla_V X$ is vertical, so

$$\nabla_V X = \mathscr{V} \nabla_V X = T_V X.$$

All the terms in the Koszul formula for $2\langle \nabla_X V, W\rangle$ vanish except one, $X(f^2(V, W))$, which is equal to $2fXf(V, W) = 2(Xf/f)f^2(V, W)$. Hence $\nabla_X V = T_V X = (Xf/f)V$.

For (3) we have

$$\begin{aligned}
\langle \nabla_V W, X\rangle &= \langle T_V W, X\rangle = -f^2(W, T_V X) \\
&= -f^2(W, V)Xf/f = -f(W, V)\langle G, X\rangle.
\end{aligned}$$

Hence $\mathscr{H} \nabla_V W = T_V W = -f(W, V)G$.

(4) follows from the fact that the induced covariant derivative on fibers is given by $\mathscr{V} \nabla_V W$, together with Remark 7.1(2)—since homometries preserve covariant derivatives.

7.4 LEMMA. *Let* X, Y, Z *be vector fields on* B, *and* U, V, W *vector fields on* F. *Then*
(1) $R_{UV}W = S_{UV}W - \|G\|^2[(U, V)V - (V, W)U]$.
(2) $R_{XV}Y = -(1/f)(\nabla^2 f)(X, Y)V = -(1/f)\langle \nabla_X G, Y\rangle V$.
(3) $R_{XY}V = R_{VW}X = 0$.
(4) $R_{XV}W = R_{XW}V = f(V, W)\nabla_X G$.
(5) $R_{XY}Z$ *is the same on either* B *or* M.

Proof. (1) follows, by the Gauss equation, from the observation (b) above, once we show that $R_{UV}W$ is vertical. But $\langle R_{UV}W, X\rangle = -\langle R_{UV}X, W\rangle$, so it suffices to prove that $R_{UV}X = 0$. We can assume that $[U, V] = 0$; thus using Lemma 7.3 we get

$$R_{UV}X = -\nabla_U \nabla_V X + \nabla_V \nabla_U X = -\nabla_U((Xf/f)V) + \nabla_V((Xf/f)U).$$

But since Xf/f is constant on fibers, this reduces to $(Xf/f)(-\nabla_U V + \nabla_V U) = -(Xf/f)[U, V] = 0$.

A similar computation, starting from $[X, V] = 0$, proves (2). In particular, $R_{XV}Y$ is symmetric in X and Y; hence the cyclic symmetry of curvature gives $R_{XY}V = 0$, which completes the proof of (3). The cyclic symmetry of R, and $R_{VW}X = 0$, give $R_{XV}W = R_{XW}V$. The rest of (4) follows from

$$\begin{aligned}
\langle R_{XV}W, Y\rangle &= -\langle R_{XV}Y, W\rangle = (1/f)\langle \nabla_X G, Y\rangle\langle V, W\rangle \\
&= f(V, W)\langle \nabla_X G, Y\rangle;
\end{aligned}$$

and $\langle R_{XV}W, U\rangle = \langle R_{WU}X, V\rangle = 0$.

Finally, (5) follows since the leaves are totally geodesic and isometric to B.

Using this lemma we now compute sectional curvature. Let Π be a plane tangent to M at $m = (b, p)$, and let vectors $x + v$, $y + w$ be an orthonormal basis for Π, where x and y are horizontal and v and w are vertical. Expansion of the expression

$\langle R_{x+v,y+w}(x+v), y+w \rangle$ yields sixteen terms; by Lemma 7.4 ten are zero and the remaining six combine to yield the *sectional curvature formula*:

$$K(\Pi) = K(x,y)\|x \wedge y\|^2$$
$$-f(b)\{(w,w)\nabla^2 f(x,x) - 2(v,w)\nabla^2 f(x,y) + (v,v)\nabla^2 f(y,y)\}$$
$$+f^2(b)[L(v,w) - \|G(b)\|^2](v \wedge w, v \wedge w).$$

The expression within braces has the following interpretation: the symmetric bilinear forms (,) on F_p and $\nabla^2 f$ on B_b combine to yield a symmetric bilinear form (,) on the direct sum $M_m = B_b + F_p$. Extending (,) to bivectors in the usual way, we find that the expression within braces is the square norm of $x \wedge w - y \wedge v$. If $\nabla^2 f$, and hence (,), is positive [semi-]definite, this norm is positive [nonnegative].

7.5 THEOREM. *Let B and F be riemannian manifolds, and let $f > 0$ be a differentiable function on B. Then the warped produce $M = B \times_f F$ has curvature $K < 0$ if the following conditions hold*:

(1) dim $B = 1$, *or $K < 0$ on B.*

(2) *f is strictly convex.*

(3) (a) dim $F = 1$, *or* (b) $L < 0$ *if f has a minimum; $L \leq 0$ if f does not have a minimum. (L is the sectional curvature of F.)*

Proof. If (1), (2), and (3) hold, it is clear that all three parts of the formula above for $K(\Pi)$ are nonpositive. If x and y are linearly independent, the first term is negative; if v and w are linearly independent, the last term is negative. If otherwise, then $(x+v) \wedge (y+w) = x \wedge w - y \wedge v \neq 0$, and the expression in braces is positive. Hence $K < 0$.

Usually the converse is true:

7.6 LEMMA. *If B is complete, and $M = B \times_f F$ has negative curvature, then conditions (1), (2), and (3) of the preceding theorem hold.*

Proof. If dim $B > 1$, then for any orthonormal x, y in B_b choosing $v = w = 0$ gives $K(\Pi) = K(x,y) < 0$. In any case, choosing $y = 0$, $v = 0$ gives $K(\Pi) = -f(b)\nabla^2 f(x,x) < 0$, so f is strictly convex.

We use the completeness of B to prove that inf $\|G\| = 0$. For otherwise $\|G\| \geq c > 0$, so if α is an integral curve of $-G/\|G\|$, then $(f \circ \alpha)' = -\|G\| \leq -c$. But the unit vector field $-G/\|G\|$ must be complete, and hence for t sufficiently large we get $f(\alpha(t)) < 0$, contradicting $f > 0$.

Now choosing $v, w \in F_p$, if $L(v,w) > 0$, we could choose b so that $\|G(b)\|^2 < L(v,w)$; and letting $m = (b,p)$, $x = y = 0$, we would have $K(\Pi) > 0$.

7.7 REMARK. Similarly: if B is complete, then $M = B \times_f F$ has $K \leq 0$ if and only if (1) dim $B = 1$ or $K \leq 0$ on B, (2) f is convex, and (3) dim $F = 1$ or $L \leq 0$.

Using the following one can construct convex functions on manifolds that have some positive curvature.

7.8 LEMMA. *Let $f > 0$ be a convex function on B. Then on $M = B \times_f F$ we have*

(1) *The horizontal and vertical subspace of M are orthogonal with respect to the bilinear form $\nabla^2 f$ on M.*

(2) *$\nabla^2 f$ is positive semidefinite on horizontal subspaces, and is positive definite whenever $\nabla^2 f | B$ is.*

(3) *$\nabla^2 f$ is positive definite on vertical subspaces at which f does not assume a minimum; $\nabla^2 f$ vanishes identically on vertical subspaces at minimum points.*

(4) *f is convex on M.*

(5) *f is strictly convex on M if and only if f is strictly convex on B and has no minimum.*

(6) *f has a minimum on M if and only if f has a minimum on B.*

Proof. If X is a vector field on B, V a vector field on F, then, since $\nabla_X G$ is horizontal, we have $\nabla^2 f(X, V) = \langle \nabla_X G, V \rangle = 0$, which proves (1). Since the restriction of $\nabla^2 f$ to horizontal subspaces coincides, essentially, with $\nabla^2 f | B$, (2) is trivial. For V as above, we have by Lemma 7.3 that $\nabla^2 f(V, V) = -\mathscr{H}(\nabla_V V) f = f(V, V) \|G\|^2 \geqq 0$. This implies (3), and the remaining assertions are clear.

7.9 REMARKS. (1) More generally, a convex function h on B lifts to a convex function on $M = B \times_f F$ if and only if $\langle G, \operatorname{grad} h \rangle \geqq 0$.

(2) Using results from above we can construct a counterexample to a converse of Theorem 3.12, i.e. show there exists a complete $K < 0$ riemannian manifold diffeomorphic to $L \times R$ which does not admit a convex function without minimum. In fact, let L be a complete $K < 0$ manifold that contains a closed geodesic, and let $f > 0$ be a strictly convex function on R with a minimum at 0. Then $R \times_f L$ is a complete $K < 0$ manifold and has the totally geodesic submanifold $\{0\} \times L$ which is homometric to L (by Remark 7.1). Thus $R \times_f L$ has a closed geodesic. It follows that any convex function on $R \times_f L$ takes on its minimum at each point of this geodesic. (Thus also $R \times_f L$ admits no strictly convex function, with or without minimum.)

(3) If L is a complete $K \leqq 0$ riemannian manifold, then the product manifold $R \times L$ has a complete $K < 0$ metric that admits a convex function without minimum. In fact, $R \times_f L$, where $f > 0$ is convex without minimum has these properties.

A specialization of this last remark gives a way to construct negative space forms. In fact, if Π is a tangent plane to $R \times_f F$ at (t, p), then Π has orthonormal basis $x + v$, w with v, w vertical and x horizontal. The sectional curvature formula then reduces to

$$K(\Pi) = -\frac{f''(t)}{f(t)} \|x\|^2 + \frac{L(v, w) - f'^2(t)}{f^2(t)} \|v\|^2$$

where $\|x\|^2 + \|v\|^2 = 1$ (warped product norm).

7.10 COROLLARY. *If a manifold F admits a complete metric with constant curvature $C \leqq 0$, then the product manifold $R \times F$ admits a complete metric with constant negative curvature.*

In fact, by the formula above (and 7.2) $R \times_f F$ has the required properties, where $f = e^t$ if $C = 0$ and $f = \cosh t$ if $C = -1$. For example, hyperbolic d-space H^d is a warped product of R and either E^{d-1} or H^{d-1}. If we give $R^+ = (0, \infty)$ the complete metric such that $\|d/du\|^2(t) = f(t) = 1/t^2$, then f is convex and $R^+ \times_f E^{d-1}$ is the Poincaré halfspace model of H^d.

In the next section we will need the following characterization of Killing fields on a warped product.

7.11 LEMMA. *A vector field Z on $B \times_f F$ is a Killing field if and only if*

(1) $\mathscr{H} Z(\cdot, q)$ *is a Killing field on B for each $q \in F$,*

(2) $\mathscr{V} Z(b, \cdot)$ *is a conformal field on F with magnification factor $-(Zf/f)(b, \cdot)$ for each $b \in B$, and*

(3) $V\langle Z, X \rangle = -f^2 X(\mathscr{V} Z, V)$ *for all vector fields X on B and V on F.*

Proof. We show that these conditions are necessary and sufficient for A_Z to be skew-symmetric. It suffices to apply A_Z to pairs X, X and V, V and X, V, where X and V are as in (3).

(1) When X is considered as a vector field on $B \times_f F$, then ∇_X commutes with the projections $\mathscr{H}$ and $\mathscr{V}$, since the horizontal leaves are totally geodesic. Thus $\langle A_Z X, X \rangle = \langle \nabla_X Z, X \rangle = \langle \nabla_X (\mathscr{H} Z), X \rangle$ which on $\eta^{-1}(q)$ is $\langle \nabla_X \mathscr{H} Z(\cdot, q), X \rangle$. Thus A_Z is skew-symmetric on horizontal vectors if and only if each $A_{\mathscr{H} Z(\cdot, q)}$ is skew-symmetric.

(2) For $b \in B$, $\nabla_{V(b, \cdot)} Z$ depends only on $Z(b, \cdot) = \mathscr{H} Z(b, \cdot) + \mathscr{V} Z(b, \cdot)$. Since $\mathscr{V} Z(b, \cdot)$ is a vector field on F, we have by Lemma 7.3(4) that $\nabla_V(\mathscr{V} Z) = D_V(\mathscr{V} Z)$. Locally we can write $\mathscr{H} Z$ in terms of functions on $B \times F$ and a local basis of vector fields from B. Applying Lemma 7.3(2) we obtain

$$\mathscr{V} \nabla_V (\mathscr{H} Z) = [(\mathscr{H} Z)f/f]V = (Zf/f)V.$$

Thus, abbreviating $\mathscr{V} Z(b, \cdot)$ to Z^b, we get

$$\langle \nabla_V Z, V \rangle(b, \cdot) = f^2(b)(D_V(Z^b), V) + f(b)Z(b, \cdot)(V, V).$$

Now we use the identity $D_V(Z^b) = -L_{Z^b}(V) + D_{Z^b}(V)$. It follows that A_Z is skew-symmetric on vertical vectors if and only if, for every $b \in B$, the $(1, 1)$ tensor field

$$D_{Z^b} - L_{Z^b} + (Zf/f)(b, \cdot)I$$

is skew-symmetric. But this is equivalent to $L_{Z^b}(\ ,\ ) = 2\mu(\ ,\ )$, with $\mu = -(Zf/f)(b, \cdot)$.

(3) Finally, we have

$$\langle \nabla_V Z, X \rangle = V\langle Z, X \rangle - \langle Z, \nabla_V X \rangle = V\langle Z, X \rangle - (Xf/f)\langle Z, V \rangle,$$

and

$$\langle \nabla_X Z, V \rangle = X\langle Z, V \rangle - \langle Z, \nabla_X V \rangle = Xf^2(\mathscr{V} Z, V) - (Xf/f)\langle Z, V \rangle$$
$$= (Xf/f)\langle Z, V \rangle + f^2 X(\mathscr{V} Z, V).$$

But A_Z is skew-symmetric on mixed pairs if and only if the sum of the expressions above is zero, thus proving (3).

We now indicate how to generalize the notion of warped product to bundles. First we assert that, for riemannian manifolds B and F, a homomorphism $h: \pi_1(B) \to I(F)$ gives rise to a riemannian manifold $M = M(B, F, h)$ with fiber bundle structure $F \to M \xrightarrow{\pi} B$ whose structural group is $I(F)$ and whose bundle charts $U \times F \to \pi^{-1}(U)$ are isometries. To construct M, identify $\pi_1(B)$ with the deck transformation group of the simply connected covering $\beta: B_1 \to B$, and let $\pi_1(B)$ act on the riemannian product $B_1 \times F$ by the isometries $\delta(b_1, p) = (\delta(b_1), h(\delta)p)$. This action is free and properly discontinuous, so the quotient manifold $M = (B_1 \times F)/\pi_1(B)$ has a unique riemannian structure making the natural map $\nu: B_1 \times F \to M$ a riemannian covering. The function $B_1 \times F \to B: (b_1, p) \to \beta b_1$ factors through ν to give the projection $\pi: M \to B$. If $U \subset B$ is evenly covered by β, then, for each lift $U_1 \subset B_1$, ν gives a fiber-preserving isometry of $U_1 \times F$ onto $\pi^{-1}(U)$. Thus $\pi^{-1}(U)$ is identified with the riemannian product $U \times F$—uniquely up to isometries of F. It follows that π preserves the length of horizontal vectors (riemannian submersion) and the horizontal distribution is integrable (flat connection).

If $f > 0$ is a C^∞ function on B, we can obviously warp the bundle $M = M(B, F, h)$ in the same way as a product: for $x \in M_m$, define $\|x\|^2$ to be $\|\mathscr{H}x\|^2 + f^2(\pi m)\|\mathscr{V}x\|^2$. For U as above, $\pi^{-1}(U)$ is now identified with $U \times_{f|U} F$, so previous results that are local in B hold for the warped bundle $M(f)$.

8. **Mobility grade.** Recall that the mobility grade of M (complete, $K \leq 0$) is the smallest integer k for which $Q = P^k(M)$ is either mobile or immobile. In this section we first obtain upper bounds for the dimension of the isometry group $I(M)$ as a function of k and the dimension d of M. Let $n = \dim Q$, $c_i = \dim P^{i-1}(M) - \dim P^i(M) \geq 2$, and $c = \sum_1^k c_i = d - n$. By iteration, using Proposition 6.6 (1), we get

$$\dim I(M) \leq \dim I(Q) + \frac{1}{2} \sum_{i=1}^k c_i(c_i - 1).$$

In the sum, if two of the c_i's exceed 2, we may decrease the smaller (say c_j) by 1 and increase the larger (say c_h) by 1, in which case the sum is increased by

$$\tfrac{1}{2}[(c_j - 1)(c_j - 2) + (c_h + 1)c_h - c_j(c_j - 1) - c_h(c_h - 1)] = c_h - c_j + 1.$$

Thus the maximum value of the sum, for given k and c, occurs when, say, $c_1 = c_2 = \cdots = c_{k-1} = 2$ and $c_k = c - 2k + 2$. The sum is then

$$\mu(k, c) = \tfrac{1}{2}[c(c + 3 - 4k) + 4k(k - 1)].$$

8.1 PROPOSITION. *Let M (complete, $K \leq 0$) have dimension d and mobility grade k.*
(1) *If M is ultimately mobile, then*

$$\dim I(M) \leq \tfrac{1}{2}[(d - 2k)^2 + d].$$

(2) *If M is ultimately immobile, then*

$$\dim I(M) \leq \tfrac{1}{2}[(d - 2k)^2 + 3d - 4k].$$

Proof. (1) Since Q is mobile, $\dim Q \geqq 1$ and $\dim I(Q) \leqq \frac{1}{2}n(n+1) = \frac{1}{2}(d-c)$ $\cdot (d-c+1)$. Adding this to $\mu(k, c)$ we get

$$\dim I(M) \leqq \tfrac{1}{2}[d^2 + d + 4k(k-1) + 2c(c-d-2k+1)].$$

Since this expression is quadratic in c with the coefficient of c^2 positive, the maximum will occur at one of the extreme values of c, which in this case are $2k$ and $d-1$. The bound is the same for both of these, namely, $\frac{1}{2}[(d-2k)^2 + d]$.

(2) Since Q is immobile, $\dim I(Q) = 0$, so our bound on $\dim I(M)$ is $\mu(k, c)$. The range of c in this case is $2k \leqq c \leqq d$, and the maximum of $\mu(k, c)$ occurs when $c = d$; that is,

$$\dim I(M) \leqq \mu(k, d) = \tfrac{1}{2}[(d-2k)^2 + 3d - 4k].$$

(In order to attain this maximum, Q must reduce to a point, since $\dim Q = d - c = 0$.)

Our aim now is to show the existence of negative curvature manifolds of arbitrary mobility grade (Corollary 8.4). We will construct these as iterated warped products, using

8.2 THEOREM. *Let B and F be complete manifolds with $K \leqq 0$ such that* (i) *B admits a strictly convex function, and* (ii) *there is no flat factor in the deRham decomposition of the simply connected coverings B_1 and F_1.* (*In particular, $K < 0$ will assure this.*) *Then there exists a strictly convex function f on B such that $P(B \times_f F) = B \times_f P(F)$ and $\dim I(B \times_f F) = \dim I(F)$.*

Before beginning the proof we note that

(ii') If $X_1, \ldots, X_n$ are linearly independent Killing fields on B (or F), then $A_{X_1}, \ldots, A_{X_n}$ are linearly independent.

Suppose, in fact, that $\sum a_i A_{X_i} = 0$. Then for $X = \sum a_i X_i$, we have $A_X = 0$. But then X is parallel on B, so if $X \neq 0$ there will be a flat factor R^1 in the deRham decomposition of B_1.

We will need the function f in the theorem to have no invariance under the isometries of B. More precisely

8.3 LEMMA. (i') *There is a strictly convex function $f > 0$ on B such that, if $G = \text{grad } f$, then for every Killing field $X \neq 0$ on B, $f[X, G]$ is not a Killing field—not even the zero field.*

Proof. By (i) there is a strictly convex function f_0 on B. Exponentiating if necessary, we may assume $f_0 > 0$. Let $e_1, \ldots, e_n$ be a basis of B_m, $m \in B$. If $c > 0$ is so small that $\exp_m(ce_i) = m_i$ lies in a normal neighborhood of m, then every isometry of B is determined by its action on $m = m_0, m_1, \ldots, m_d$. Thus if $X \neq 0$ is a Killing field, e^{tX} moves at least one of the m_i for all small t. Now let u be a C^∞ function having compact support and isolated maxima at each m_i. Then for small t, $u \circ e^{tX}$ has isolated maxima at $e^{tX}(m_i)$. Differentiating, we see that $L_X(du) = d(Xu) \neq 0$ for every Killing field $X \neq 0$. Moreover, $u L_X(du)$ is not a Killing 1-form, since it

vanishes outside the support of u, but is not identically zero. Now we dualize, using the riemannian metric $\langle\ ,\ \rangle$ of B. If $U = \operatorname{grad} u$, the linear space

$$\mathscr{A} = \{u[X, U] \mid X \in i(B)\}$$

has the same dimension as $i(B)$ and does not meet $i(B)$ except in 0. (Note that L_X commutes with dualization, since X is Killing.)

We wish to show that we can destroy whatever symmetry f_0 may have by adding tu; for small t, the function $f_0 + tu$ will be convex and positive (since u has compact support). Note that

$$(f_0 + tu)[X, G_0 + tU] = f_0[X, G_0] + t(u[X, G_0] + f_0[X, U]) + t^2 u[X, U].$$

It follows that all of these vector fields are in the linear space $\mathscr{B} \supset \mathscr{A}$ spanned by

$$\{f_0[X, G_0],\ u[X, G_0] + f_0[X, U],\ u[X, U],\ X \mid X \in i(B)\}.$$

Define a family of linear transformation $T_t : i(B) \to \mathscr{B}/i(B)$ by

$$T_t(X) = (f_0 + tu)[X, G_0 + tU] + i(B) = T(X) + tT'(X) + t^2 T''(X).$$

We wish to show that there exist small t for which T_t is one-one; that is, no $(f_0 + tu)[X, G_0 + tU]$ is Killing for $X \neq 0$. Choose bases for $i(B)$ and $\mathscr{B}/i(B)$ that are related by T''; that is, if the basis of $i(B)$ is $X_1, \ldots, X_n$, then that of $\mathscr{B}/i(B)$ starts with $u[X_1, U] + i(B), \ldots, u[X_n, U] + i(B)$. The latter are linearly independent, since $\mathscr{A} + i(B)$ is a direct sum. With respect to such a basis the matrix of T'' is

$$\begin{pmatrix} I \\ 0 \end{pmatrix}$$

where I is the $n \times n$ identity matrix. Thus the first minor of T_t is

$$\det (t^2 I + tA + B) = t^{2n} + \text{lower degree terms}.$$

Thus T_t has rank n, except for finitely many t. This completes the proof of Lemma 8.3.

NOTE.. In particular, the conditions on f forbid $Xf = 0$ for $0 \neq X \in i(B)$. In fact, if $Xf = 0$, then $f[X, G] = fL_X G$ is dual to $fL_X(df) = fd(Xf) = 0$.

Proof of Theorem 8.2. Let Z be a Killing field on $B \times_f F$, where f is given by the preceding lemma. By Lemma 7.11 (1) we may write $Z = \sum h_i X_i + Y$, where $Y = \mathscr{V} Z$, $h_i : F \to R$, and $X_1, \ldots, X_n$ is a basis for $i(B)$. Then by Lemma 7.11(2), for every point $b \in B$, the vector field $Y(b, \cdot)$ on F is conformal, with magnification factor $-\sum (X_i(b)f/f(b))h_i$. Since the functions $X_i f$ are linearly independent (by the Note above), we can choose points b_j $(j = 1, \ldots, n)$ such that the matrix $(X_i(b_j)f)$ is nonsingular. In fact, by a change of basis of $i(B)$ we can arrange that this matrix is the identity matrix. Thus $Y(b_j, \cdot) = Y_j$ is conformal on F with magnification factor $-\sum \delta_{ij} h_i = -h_j$.

Reordering if necessary, we may also suppose that $h_1, \ldots, h_s$ are a basis for the space spanned by the h_i over the reals. (If all the Y_j are Killing we have all $h_i = 0$

and $s=0$.) Then $h_{s+i}=\sum_1^s a_{ij}h_j$ for some constants a_{ij}, and since $Y_{s+i}-\sum a_{ij}Y_j$ has magnification factor 0, it is Killing. Similarly, $Y(b, \cdot)-\sum_1^n (X_i(b)f/f(b))Y_i$ is Killing for every $b \in B$. Letting V_α, $\alpha=1, \ldots, p$ be a basis for $i(F)$, we have

$$Y_{s+i} = \sum_{j=1}^s a_{ij}Y_j + \sum_{\alpha=1}^p b_{i\alpha}V_\alpha,$$

$$
\begin{aligned}
Y(b, \cdot) &= \sum_{i=1}^n \frac{X_i(b)f}{f(b)} Y_i + \sum_{\alpha=1}^p k'_\alpha(b)V_\alpha \\
&= \sum_{i=1}^s \frac{X_i(b)f}{f(b)} Y_i + \sum_{i=1}^{n-s} \frac{X_{s+i}(b)f}{f(b)} \left(\sum_{j=1}^s a_{ij}Y_j + \sum_{\alpha=1}^p b_{i\alpha}V_\alpha \right) + k'_\alpha(b)V_\alpha \\
&= \frac{1}{f(b)} \sum_{j=1}^s \left(X_j(b) + \sum_{i=1}^{n-s} a_{ij}X_{s+i}(b) \right) fY_j + \sum_{\alpha=1}^p k_\alpha(b)V_\alpha \\
&= \sum_{j=1}^s \frac{W_j(b)}{f(b)} Y_j + \sum_{\alpha=1}^p k_\alpha(b)V_\alpha,
\end{aligned}
$$

where $W_j = X_j + \sum_{i=1}^{n-s} a_{ij}X_{s+i}$. Then

$$
\begin{aligned}
Z &= \sum_{j=1}^s h_j X_j + \sum_{i=1}^{n-s}\sum_{j=1}^s a_{ij}h_j X_{s+i} + \sum_{j=1}^s (X_jf/f)Y_j + \sum_{\alpha=1}^p k_\alpha V_\alpha \\
&= \sum_{j=1}^s (h_j W_j + (W_jf/f) Y_j) + \sum_{\alpha=1}^p k_\alpha V_\alpha.
\end{aligned}
$$

Here the functions $k_\alpha: B \to R$ and $h_j: F \to R$ are clearly C^∞.

In this final expression we have that the magnification factors $-h_j$ of the Y_j are linearly independent, as well as the $W_j \in i(B)$. This summarizes the content of Lemma 7.11 (1)(2) in terms of bases. We have yet to apply (3), which now gives, for vector fields X on B and V on F:

$$
\begin{aligned}
V\langle Z, X \rangle &= \sum_1^s (Vh_j)\langle W_j, X \rangle \\
&= -f^2 X(Z, V) \\
&= -f^2 \left[\sum_1^s (XW_jf/f)(Y_j, V) + \sum (Xk_\alpha)(V_\alpha, V) \right].
\end{aligned}
$$

Let ω_j be the 1-form $\langle W_j, \cdot \rangle$ on B, η_j the 1-form $(Y_j, \cdot)$ and ϑ_α the 1-form $(V_\alpha, \cdot)$ on F. We identify these 1-forms with their pull-backs to $B \times F$ via the projections onto B and F. Then the formula equivalent to 7.11(3) becomes

(1) $$\sum \omega_j \wedge dh_j = -f \sum d(W_jf) \wedge \eta_j - f^2 \sum dk_\alpha \wedge \vartheta_\alpha.$$

Taking exterior derivatives yields

(2) $$
\begin{aligned}
\sum d\omega_j \wedge dh_j = &-df \wedge \sum d(W_jf) \wedge \eta_j + f \sum d(W_jf) \wedge d\eta_j \\
&-2f\, df \wedge \sum dk_\alpha \wedge \vartheta_\alpha + f^2 \sum dk_\alpha \wedge d\vartheta_\alpha.
\end{aligned}
$$

Under the (pointwise) identification of $\bigwedge^3 T^*(B \times F)$ with the direct sum

$$\bigwedge^3 T^*(B) + \bigwedge^2 T^*(B) \otimes T^*(F) + T^*(B) \otimes \bigwedge^2 T^*(F) + \bigwedge^3 T^*(F),$$

only the inner types occur in (2). Thus we have just two direct summand equations:

(3) $$\sum d\omega_j \wedge dh_j = -df \wedge \sum d(W_j f) \wedge \eta_j - 2f \, df \wedge \sum dk_\alpha \wedge \vartheta_\alpha.$$

(4) $$0 = \sum d(W_j f) \wedge d\eta_j + f \sum dk_\alpha \wedge d\vartheta_\alpha.$$

Taking the exterior derivative of (4) we get

(5) $$0 = df \wedge \sum dk_\alpha \wedge d\vartheta_\alpha.$$

Note that $d\vartheta_\alpha = \nabla \vartheta_\alpha$ since ϑ_α is a Killing 1-form and $d\vartheta$ is always the skew-symmetric part of $\nabla \vartheta$. Hence by (ii′) the forms $d\vartheta_\alpha$ are linearly independent. Thus we must have $df \wedge dk_\alpha = 0$ for each α. (The second $\wedge$ is, in effect, a tensor product on two independent spaces.) This simplifies (3) to

(6) $$\sum d\omega_j \wedge dh_j = -\sum df \wedge d(W_j f) \wedge \eta_j.$$

Again by (ii′) the $d\omega_j$ are linearly independent. More specifically, we can find s vector pairs (x_i, x_i') at various points of B such that the matrix $(d\omega_j(x_i, x_i'))$ is nonsingular. Thus (4) says, by multiplying by the inverse matrix, that the dh_j are linearly dependent on the η_j:

(7) $$dh_i = \sum c_{ij} \eta_j, \qquad c_{ij} \text{ constants.}$$

Inserting (7) in (1) we have

$$\sum_j \left(\sum_i c_{ij} \omega_i + f \, d(W_j f) \right) \wedge \eta_j + f^2 \sum dk_\alpha \wedge \vartheta_\alpha = 0.$$

The η_j and ϑ_α are linearly independent, so $dk_\alpha = 0$, that is, the k_α are constant, and

(8) $$\sum c_{ij} \omega_i + f \, d(W_j f) = 0.$$

But $f \, d(W_j f) = f L_{W_j}(df)$ corresponds to $f[W_j, G]$ under duality by $\langle \, , \, \rangle$. Thus (8) contradicts Lemma 8.3. The only relief from this contradiction is the possibility $s = 0$; that is

$$Z = \sum k_\alpha V_\alpha, \qquad k_\alpha \text{ constants.}$$

Conversely, we have seen that the injection of $i(F)$ into the vector fields of $B \times_f F$ produces Killing fields. The conclusions of the theorem are now clear.

   A point p of a riemannian manifold M is called a *central point* of M provided (1) every isometry of M leaves p fixed, and (2) every linear isometry of M_p into

itself is the differential map of an isometry of M. We can furnish the manifold R^d with riemannian structures such that the origin 0 is a central point by the following warping of polar coordinates. Represent $N = R^d - \{0\}$ as a product manifold of $R^+ = (0, \infty)$ and the unit sphere S^{d-1} by means of the projections $\pi(n) = \|n\|$ and $\eta(n) = n/\|n\|$. If $f: R^+ \to R$ is a function possessing a C^∞ extension to R that is odd and has $f'(0) = 1$, then the warped product structure on N and the usual euclidean inner product on $(R^d)_0$ combine to give a C^∞ riemannian structure on R^d. Obviously condition (2) above holds at $p = 0$ for the resulting (necessarily complete) riemannian manifold $R^d(f)$. The curvature formula preceding Corollary 7.10 now applies with $L = 1$. Thus, for example, $R^d(\sinh)$ is hyperbolic space. If we choose $f(t) = t + t^3/6$, then $R^d(f)$ has curvature $-1 \le K < 0$, with $K(\Pi) = -1$ if and only if Π is a plane at the origin. Thus 0 is clearly a central point.

Conversely, if M (complete, $K < 0$) has a central point p, then $P(M) = \{p\}$, since p is an isolated point of the connected set $P(M)$. Thus M is simply connected and ultimately immobile. Furthermore, M is isometric to $R^d(f)$ for suitable f.

We can now give the existence proof mentioned earlier.

8.4 COROLLARY. *Given integers $d \ge 2$ and $c_1, \ldots, c_k$ satisfying the necessary conditions $c_i \ge 2$ and $\sum c_i \le d$, there exists a complete d-dimensional riemannian manifold M with $K < 0$ such that*

(1) *M has mobility grade k, and*

(2) *$P^i(M)$ has codimension c_i in $P^{i-1}(M)$.*

Furthermore, if $d - \sum c_i \ge 2$, then M can be required to be ultimately mobile or to be ultimately immobile.

(In the last assertion, $n = d - \sum c_i$ is the dimension of $P^k(M)$; hence M must be ultimately mobile if $n = 1$, ultimately immobile if $n = 0$.)

Proof. First we assert that for $d \ge 2$ there exists a complete $K < 0$ manifold M of dimension d and such that M admits a strictly convex function without minimum. Furthermore, such M exist that are (a) mobile and (b) immobile. For (a) we can use the hyperbolic space $H^d = R \times_{e^t} R^{d-1}$ with f the function e^t (lifted as usual to H^d). For (b) a small perturbation of the riemannian structure of H^d will produce an immobile $K < 0$ manifold that retains the other properties of H^d and f. (This remark includes the $k = 0$ case of the corollary.)

Now for $k > 0$ consider the case where $n = d - \sum c_i$ is at least 2. Let P_k be a complete, $K < 0$, n-dimensional manifold admitting a strictly convex function without minimum. Let F be complete, $K < 0$, c_k-dimensional with a central point q. Then by Theorem 8.2 there is a strictly convex function f on P_k such that if $P_{k-1} = P_k \times_f F$ we have $P(P_{k-1}) = P_k \times \{q\}$. By results of the preceding section, P_{k-1} is again complete, $K < 0$ and admits a strictly convex function without minimum. By iteration we reach the required manifold M. The final assertion holds, since by the initial assertion of this proof, we can choose P_k to be mobile or to be immobile.

If $n = 0$, we start the iteration with P_{k-1} a complete $K < 0$, c_k-dimensional

manifold with a central point (which thus becomes $P^k(M)$). If $n = 1$, we start with P_{k-1} a $(c_k + 1)$-dimensional manifold of the type given by the following lemma. In both of these cases, P_{k-1} admits a strictly convex function without minimum. In fact, P_{k-1} is simply connected, so the required function can be constructed by applying Theorem 4.1 (3) with S and T asymptotic geodesics (see §9).

8.5 LEMMA. *Let F be a complete manifold with sectional curvature $L < 0$ and with a central point q. Then $N = R \times_{e^t} F$ is complete, $K \leq -1$, with*

$$I(N) = \{1 \times \psi \mid \psi \in I(F)\}.$$

In particular, $P(N) = R \times \{q\}$.

Proof. From the curvature formula preceding Corollary 7.10 we compute $K(\Pi) = -1 + e^{-2t}L(v, w)\|v\|^2$, where $x + v$, w is the usual orthonormal basis for Π at $(t, p) \in N$. Since L is never 0, $K(\Pi) = -1$ if and only if $v = 0$; that is, if and only if Π contains a nonzero horizontal vector.

Let $\phi \in I(M)$. If $x \neq 0$ is a horizontal vector, then every plane through x has $K = -1$; hence every plane through $\phi_*(x)$ has $K = -1$; hence $\phi_*(x)$ is horizontal. Thus ϕ carries leaves to leaves. But then ϕ carries vertical vectors to vertical vectors, hence fibers to fibers. Thus ϕ is a product function $\phi_1 \times \phi_2$. It is clear that ϕ_1 is an isometry, say $\phi_1(t) = \pm t + a$. For any $t \in R$, we can express ϕ_2 as the composition of: the injection of F onto $\pi^{-1}(t)$, the isometry ϕ from $\pi^{-1}(t)$ to $\pi^{-1}(\pm t + a)$, and the projection η of $\pi^{-1}(\pm t + a)$ onto F. By Remark 7.1 (2) it follows that ϕ_2 is a homometry with scale factor either e^{-a} or e^{2t-a}. The latter is impossible, since this number must not depend on t. But a selfhomometry of a complete nonflat riemannian manifold is an isometry (cf. [6], p. 242). Hence $a = 0$. Thus $\phi = 1 \times \phi_2$, with $\phi_2 \in I(F)$. Conversely, if $\psi \in I(F)$, then $1 \times \psi \in I(N)$. The final assertion follows since $P(F) = \{q\}$.

9. Asymptotes. We first review briefly the basic facts about asymptotes; compare Buseman [2].

Let M be a complete, simply connected riemannian manifold. We say that a geodesic ray $\rho: [0, \infty) \to M$ is *asymptotic* to a geodesic ray σ provided there is a sequence $\{\rho_i\}$ of geodesic rays and sequences $\{s_i\}$, $\{t_i\}$ of positive numbers such that (a) $\lim \rho_i'(0) = \rho'(0)$, (b) $\rho_i(s_i) = \sigma(t_i)$ for all i, and (c) $\lim t_i = \infty$. (Clearly this definition does not depend in any essential way on the parametrizations of ρ and σ.) If M is not simply connected, we say that ρ is *asymptotic* to σ provided they have lifts ρ_1, σ_1 to the simply connected covering M_1 such that ρ_1 is asymptotic to σ_1.

9.1 LEMMA. *For any geodesic ray σ and point $m \in M$, there is a geodesic ray ρ starting at m and asymptotic to σ.*

Proof. Evidently we can assume that M is simply connected. Let ρ_i be a geodesic ray starting at m and passing through $\sigma(i)$. Then some subsequence of $\{\rho_i'(0)\}$ will converge to the required $\rho'(0)$.

9.2 PROPOSITION. *Let ρ and σ be geodesic rays in a complete, simply connected manifold H with $K \leqq 0$. Then ρ is asymptotic to σ if and only if the function $s \to d(\rho(s), \sigma)$ is bounded.*

Proof. Suppose ρ is asymptotic to σ, and as in the definition let ρ_i be an approximating sequence with $\rho_i(s_i) = \sigma(t_i)$ for all i. Since $\sigma([0, \infty))$ is a closed, totally convex set, it follows by Proposition 4.7 that $f = d(\cdot, \sigma)$ is a continuous convex function. Fix $s \geqq 0$. By definition, $\lim t_i = \infty$; it follows that $\lim s_i = \infty$. Let i be large enough so that $s_i \geqq s$. Now $f \circ \rho_i$ is continuous and convex, and $f(\rho_i(s_i)) = f(\sigma(t_i)) = 0$, hence $f(\rho_i(s)) \leqq f(\rho_i(0))$. But $\lim \rho_i'(0) = \rho'(0)$ implies $\lim \rho_i(t) = \rho(t)$ for all t. Thus taking limits in the preceding inequality gives $f(\rho(s)) \leqq f(\rho(0))$; hence $d(\rho(s), \sigma)$ is bounded.

Conversely, suppose b is an upper bound for $d(\rho(s), \sigma)$, and let $\{u_i\}$ be a positive sequence such that $\lim u_i = \infty$. Let $\sigma(t_i)$ be the foot (that is, nearest point) of $\rho(u_i)$ on σ. Assuming for simplicity that these rays have unit speed, we find, using the triangle inequality, that

$$t_i = d(\sigma(0), \sigma(t_i)) \geqq u_i - d(\rho(0), \sigma(0)) - d(\rho(u_i), \sigma(t_i)).$$

Since the last term is bounded by b, $\lim t_i = \infty$. Let ρ_i be the (unit speed) geodesic ray starting at $\rho(0)$ and passing through $\sigma(t_i)$ with $\rho_i(s_i) = \sigma(t_i)$. It remains to show that $\lim \rho_i'(0) = \rho'(0)$, or, equivalently, that $\lim \cos \vartheta_i = 1$, where ϑ_i is the angle between $\rho'(0)$ and $\rho_i'(0)$. But from the geodesic triangle $\rho(0)$, $\rho(u_i)$, $\rho_i(s_i)$ we get

$$b^2 \geqq d_i^2 \geqq u_i^2 + s_i^2 - 2u_i s_i \cos \vartheta_i \geqq 2u_i s_i (1 - \cos \vartheta_i),$$

where $d_i = d(\rho(u_i), \sigma(t_i)) = d(\rho(u_i), \sigma)$. The result follows.

It is clear from this proof that in the definition of asymptote there is no loss of generality in assuming that each ρ_i starts at the same point as ρ.

9.3 COROLLARY. *Geodesic rays ρ and σ in H having the same speed are asymptotic if and only if $d(\rho(t), \sigma(t))$ is bounded.*

Proof. Fix the notations: $d_0 = d(\rho(0), \sigma(0))$, and $\sigma(t_1)$ the foot of $\rho(t)$ on σ. Suppose ρ is asymptotic to σ (both unit speed). Following Busemann [2, p. 250] we get

$$d(\sigma(t), \sigma(t_1)) = |t - t_1| = |d(\rho(0), \rho(t)) - d(\sigma(0), \sigma(t_1))| \leqq d_0 + d(\rho(t), \sigma(t_1)).$$

But by Proposition 9.2 this last term is bounded, say by $b > 0$. Hence

$$d(\rho(t), \sigma(t)) \leqq d(\rho(t), \sigma(t_1)) + d(\sigma(t_1), \sigma(s)) \leqq d_0 + 2b.$$

The converse implication is immediate from Proposition 9.2.

9.4 COROLLARY. *Asymptoticness is an equivalence relation on the set of geodesic rays in H.*

If asymptotic geodesic rays meet in H (simply connected) it is easy to see that they are the same but for parametrization. Thus the preceding corollary gives

9.5 COROLLARY. *If σ is a geodesic ray in H, then the geodesic ray starting at $m \in H$ and asymptotic to σ is unique.*

9.6 PROPOSITION. *Let σ be a geodesic ray in H. For each point $p \in H$, let $V(p)$ be the initial velocity of the unit speed geodesic from p asymptotic to σ. Then V is a continuous vector field.*

Proof. For each $i \geq 1$, let V_i be the C^∞ vector field on $H - \sigma(i)$ such that $V_i(p)$ is the initial velocity of the unit speed geodesic ρ_i from p to $\sigma(i)$. We know that $\lim V_i(p) = V(p)$ for each point p; now we prove that this convergence is uniform on compact sets. Let B be a compact set containing $p \in H$. Let $b > 0$ be an upper bound for the distances from points of B to σ. It suffices to show that the angle ϑ_{ij} between $V_i(p)$ and $V_j(p)$ approaches zero uniformly (in $p \in B$) as i and j go to infinity. Let q be the foot of p on σ. Thus $b \geq d(p, q)$. Then in the geodesic triangle $p\sigma(i)q$, it is easy to see that the angle α_i at $\sigma(i)$ uniformly approaches zero. If $i < j$, then for the triangle p, $\sigma(i)$, $\sigma(j)$ we have angle sum $\vartheta_{ij} + (\pi - \alpha_i) + \alpha_j \leq \pi$. Thus $\vartheta_{ij} \leq \alpha_i - \alpha_j$, and the result follows.

For a geodesic $\gamma: R \to M$ call $\gamma^+ = \gamma|[0, \infty)$ the *positive ray* of γ, while the *negative ray* γ^- of γ is defined by $\gamma^-(u) = \gamma(-u)$ for $u \geq 0$. Then two geodesics are said to be *asymptotic* provided their positive rays are asymptotic.

From now on we do not assume simple connectedness, unless explicitly mentioned. A routine check shows that asymptoticness is still an equivalence relation—on entire geodesics as well as geodesic rays. However, the criteria in 9.2 and 9.3 fail in general, and it is possible for distinct geodesics through the same point to be asymptotic.

Let f be a convex function on M. An *f-monotone* geodesic is a geodesic on which f is monotone nonincreasing. The rest of this section deals with continuous as well as C^∞ convex functions.

9.7 PROPOSITION. *Let f be convex without minimum on a complete manifold M. For each point $m \in M$ there is an f-monotone geodesic starting at m such that the positive ray γ^+ is minimal.*

Proof. Let $\{a_i\}$ be a sequence such that $a_i < f(m)$ and $\lim a_i = \inf f$. For each i, let γ_i be a unit speed geodesic that is minimal from m to $f^{-1}(a_i)$. Passing if necessary to a subsequence, we have $\lim \gamma_i'(0) = x \in M_m$. We assert that the geodesic γ with initial velocity x has the required properties. Let s_i be the first number for which $\gamma_i(s_i) \in f^{-1}(a_i)$. Clearly $\lim s_i = \infty$. Fix $s \geq 0$. By construction, $f(\gamma_i(0)) = f(m) > a_i = f(\gamma_i(s_i))$. If i is large enough, then $0 \leq s \leq s_i$, and since f is convex we obtain $f(\gamma_i(s)) \leq f(m)$. But $\lim \gamma_i(s) = \gamma(s)$, hence $f(\gamma(s)) \leq f(m)$. Again since f is convex, this boundedness implies that γ is f-monotone. Furthermore, by construction $s = d(\gamma_i(0), \gamma_i(s))$. Taking limits gives $s = d(\gamma(0), \gamma(s))$; thus γ^+ is minimal.

9.8 PROPOSITION. *Let f be a convex function on a complete manifold M. A geodesic asymptotic to an f-monotone geodesic is f-monotone.*

Proof. It suffices to assume M is simply connected. Let ρ be asymptotic to the f-monotone geodesic σ, and let ρ_i be an approximating sequence with $\rho_i(s_i) = \sigma(t_i)$. Fix $s \geq 0$. Since $f \circ \rho_i$ is convex, for $s_i \geq s$ we have $f(\rho_i(s)) \leq \max \{f(\rho_i(0)), f(\rho_i(s_i))\}$. But $f(\rho_i(s_i)) = f(\sigma(t_i)) \leq f(\sigma(0))$; hence in the limit we obtain

$$f(\rho(s)) \leq \max \{f(\rho(0)), f(\sigma(0))\}.$$

Since $f \circ \rho$ is convex it follows that ρ is f-monotone.

9.9 REMARK. The f-monotone geodesic through a point need not be unique, even if f is strictly convex and M has curvature $K \leq 0$. We give two examples.

(1) $K = 0$. Let M be the euclidean plane with cartesian coordinates x, y. Let g be a strictly convex function on the real line. Then the function $f(x, y) = g(x) + g(y)$ is strictly convex on M. However, if we choose g without minimum, then from each point there will be a quadrant of directions for f-monotone geodesics. If we choose g bounded below, this example shows that there can be many values for

$$\lim_{t \to \infty} f(\gamma(t))$$

as γ runs through the f-monotone geodesics.

(2) $K < 0$, $\sup K = 0$. Let M be the surface of revolution $z = f(r)$, where $r = (x^2 + y^2)^{1/2}$. Then the curvature $K = f'f''/r(1 + f'^2)^2$ is negative if f is decreasing and strictly convex. (In fact, M is the warped product $R \times_g S^1$, where $r = g(z)$ is the inverse of f.) If we choose g bounded away from 0 by $c > 0$ and so that $\lim_{r \to c} f(r) = +\infty$, then the r-monotone geodesic through a point is not unique. Indeed, a geodesic is r-monotone if it continues up the tube without doubling back. Let α be the angle a geodesic makes with the vertical geodesic. It is well known that $r \sin \alpha$ is constant on the geodesic. By choosing the initial value of α so that $r \sin \alpha \in [-c, c]$, we prevent α from ever becoming $\pi/2$, so the geodesic must be r-monotone. Thus we can get infinitely many of these from any given point.

Our aim now is to show that uniqueness does obtain when M has strictly negative curvature, that is, $K \leq c < 0$.

9.10 LEMMA. *Let ρ and σ be distinct geodesic rays starting at $m \in M$ (complete, $K \leq c < 0$). There is a geodesic γ such that γ^+ is asymptotic to σ and γ^- is asymptotic to ρ.*

Proof. We can suppose that M is simply connected, and ρ and σ have unit speed. For $i \geq 1$, let γ_i be the unit speed geodesic through $\rho(i)$ and $\sigma(i)$. Since $d^2(m, \cdot)$ is strictly convex, the foot $\gamma_i(t_i)$ of m on γ_i lies between $\rho(i)$ and $\sigma(i)$. Let $d_i = d(m, \gamma_i(t_i))$. Let A_i be the area of the triangle composed of all geodesic segments from m to points of γ_i between $\rho(i)$ and $\sigma(i)$. On each such radial segment the points with distance at most d_i from m form a patch of surface that standard estimates show to have area greater than that of a euclidean sector of angle $\vartheta = \sphericalangle(\rho'(0), \sigma'(0))$ and radius d_i. Thus $A_i > \vartheta d_i^2/2$. On the other hand, the Gauss-Bonnet theorem implies that $A_i < \pi/|c|$, where $K \leq c < 0$. Hence d_i remains bounded

as i goes to infinity. Thus some subsequence of $\gamma_i'(t_i)$ converges to a tangent vector x and the geodesic γ with initial velocity x has the required properties.

9.11 PROPOSITION. *Let M be complete and have strictly negative curvature. If f is convex on M, has no minimum, and is not constant on any geodesic, then the f-monotone geodesic starting at a point of M is unique. In particular this obtains if f is strictly convex without minimum.*

Proof. Again it suffices to prove the simply connected case. Assume ρ and σ are distinct f-monotone geodesics starting at $m \in M$. Let γ be as in the preceding lemma. Then by Proposition 9.8, f is monotone on both γ^+ and γ^-, so f is constant on γ, a contradiction.

By an *asymptote class* we mean an equivalence class of mutually asymptotic geodesics. By Proposition 9.8, f-monotonicity is a property of asymptote classes.

9.12 THEOREM. *If M (complete, $K \leq c < 0$) admits a strictly convex function and is not simply connected, then M contains a unique asymptote class that is f-monotone for every convex function f on M. (We call this the principal asymptote class of M.)*

Proof. Let g be strictly convex on M. Since M is not simply connected, g has no minimum. We show first that the set A of all g-monotone geodesics constitutes exactly one asymptote class. By Proposition 9.7, A is not empty. If $\rho, \sigma \in A$, let τ be an asymptote to ρ from a point of σ. Then τ is g-monotone by Proposition 9.8; hence $\sigma = \tau$ by Proposition 9.9. Thus ρ and σ are asymptotic. By Proposition 9.8. again, A is exactly one asymptote class.

To show that A is principal, let f be a convex function on M, and assume that A is *not* f-monotone. Then $\lim f \circ \gamma = \infty$ for one (hence every) $\gamma \in A$. Now consider $h = e^f + e^g$. If $\gamma \in A$, then $\lim h \circ \gamma = \infty$ since $\lim f \circ \gamma = \infty$. If $\gamma \notin A$, then $\lim h \circ \gamma = \infty$, by Proposition 9.11. Thus no geodesic is h-monotone, contradicting Proposition 9.7.

Nonsimple connectedness is necessary in the preceding theorem, for if M is simply connected we have seen that $d^2(m, \cdot)$ is strictly convex with minimum; thus M cannot have a principal asymptote class.

9.13 REMARKS. We give some properties of the principal asymptote class of a manifold M as in Theorem 9.12.

(1) Each point of M is on a unique principal asymptote.

(2) Principal asymptotes minimize arc length. (This follows from the minimality property in Proposition 9.7.)

(3) If $\pi: M_1 \to M$ is the simply connected covering, the lifts to M_1 of the principal asymptotes of M constitute a single asymptote class. In fact, if g is strictly convex on M (hence without minimum), then $g \circ \pi$ is strictly convex without minimum on M_1. But the $(g \circ \pi)$-monotone class consists of the lifts of all g-monotone geodesics, and the latter form the principal asymptote class of M.

The following proof uses two properties of a Killing field X: If γ is a geodesic,

then $\langle X, \gamma' \rangle$ is constant on γ. If α is an integral curve of X, then $\|X\|$ is constant on α. (Both properties follow at once from the skew-symmetry of A_X.)

9.14 PROPOSITION. *For M as in Theorem* 9.12, (1) *the principal asymptote class is invariant under all isometries of M, and* (2) *every Killing field X on M is orthogonal to every principal asymptote.*

Proof. (1) If φ is an isometry and f is strictly convex, then any principal asymptote γ is f-monotone. Hence $\varphi \circ \gamma$ is $(f \circ \varphi^{-1})$-monotone. Then by Theorem 9.12, $\varphi \circ \gamma$ is a principal asymptote.

(2) Let $f = \|X\|$, so f is convex. By Proposition 5.5 there are three cases: (a) X is zero at some point $p \in M$, (b) X has a nonconstant geodesic integral curve α, (c) f is strictly convex without minimum. In Case (a), let σ be the principal asymptote through p. Since σ is f-monotone, f is zero on σ^+, hence on all of σ. Now let ρ be an arbitrary principal asymptote, and let $\{\rho_i\}$ be an approximating sequence from $m = \rho(0)$ to σ. Since ρ_i meets σ, one of the above mentioned properties of Killing fields shows that $\langle X, \rho_i' \rangle = 0$; in particular, $\langle X(m), \rho_i'(0) \rangle = 0$. Taking limits gives $\langle X(m), \rho'(0) \rangle = 0$.

Case (b) cannot occur. In fact, since M is not simply connected, Lemma 6.2 shows that α cannot be one-one, and since M admits a strictly convex function, α cannot be closed.

Finally, for Case (c), let $m \in M$, and as in the proof of Proposition 9.7 let $\{a_i\}$ be a sequence such that $a_i < f(m)$ and $\lim a_i = \inf f$. Let γ_i be a geodesic that is minimal from m to $f^{-1}(a_i)$. We have seen that for a subsequence, $\lim \gamma_i'(0) = \gamma'(0)$, where γ is f-monotone, hence principal. Suppose $\gamma_i(s_i)$ is the first point of γ_i in $f^{-1}(a_i)$. Then by minimality, $\gamma_i'(s_i)$ is orthogonal to the hypersurface $f^{-1}(a_i)$. By the other Killing field fact above, X is tangent to this hypersurface and hence orthogonal to $\gamma_i'(s_i)$. Thus $\langle X(m), \gamma_i'(0) \rangle = 0$, and so X is orthogonal to $\gamma'(0)$—and in fact to γ'.

10. Fundamental group. Let M be a complete manifold with $K \leqq 0$, and let $\gamma: [0, \infty) \to M$ be a geodesic ray with $\gamma(0) = m$. If $\alpha: I \to M$ is a geodesic loop at m, let α_t be the unique geodesic loop at $\gamma(t)$ contained in the homotopy class $[\gamma|[0, t]^{-1} \cdot \alpha \cdot \gamma|[0, t]] \in \pi_1(M, \gamma(t))$. Thus $[\alpha] \to [\alpha_t]$ is the isomorphism $\pi_1(M, m) \approx \pi_1(M, \gamma(t))$ determined by $\gamma|[0, t]$.

10.1 LEMMA. *The length function $t \to |\alpha_t|$ is C^∞ convex. If $K < 0$, this function is constant or strictly convex.*

Proof. Let $\pi: M_1 \to M$ be the simply connected covering of M. Fix m_1 such that $\pi(m_1) = m$. Let γ_1 be the lift of γ starting at m_1 and let β_t be the lift of α_t starting at $\gamma_1(t)$. Thus $r(u, t) = \beta_t(u)$ defines a C^∞ rectangle r in M_1 whose length function is $L(t) = |\alpha_t| = |\beta_t|$, which we wish to show is convex.

Let $X = \partial r / \partial u$ and $Y = \partial r / \partial t$ be the longitudinal and transverse vector fields on r.

Since the β_t are geodesic segments parametrized on I, their lengths coincide with those of X, that is, $L = \|X\|$. Note that Y is a Jacobi field along each β_t and $\nabla_X X = 0$. Now differentiating L^2 yields:

$$LL' = \langle \nabla_Y X, X \rangle = \langle \nabla_X Y, X \rangle = \frac{\partial}{\partial u} \langle Y, X \rangle.$$

Differentiating again:

$$LL'' = \frac{\partial}{\partial u} (\langle \nabla_Y Y, X \rangle + \langle Y, \nabla_Y X \rangle) - (L')^2$$

$$= \frac{\partial}{\partial u} \langle \nabla_Y Y, X \rangle + \langle Y, \nabla_X \nabla_X Y \rangle + \langle \nabla_X Y, \nabla_X Y \rangle - \langle \nabla_X Y, X \rangle^2 / \langle X, X \rangle$$

$$= \frac{\partial}{\partial u} \langle \nabla_Y Y, X \rangle - K \langle X, Y \rangle \|X \wedge Y\|^2 + \|X \wedge \nabla_X Y\|^2 / L^2.$$

The terms in this formula may individually depend on u, but the sum is a function of t only. Because $r(0, \cdot)$ and $r(1, \cdot)$ are both lifts of the geodesic γ, $\nabla_Y Y = 0$ when $u = 0$ or 1. Hence by Rolle's theorem there exists $h(t)$ such that $(\partial/\partial u)\langle \nabla_Y Y, X \rangle = 0$ when $u = h(t)$. Thus $L'' \geq 0$ if L is nonzero. However, if L ever vanishes, then α is the constant loop, so L always vanishes and L is certainly convex.

Now suppose $K < 0$, α is nonconstant, and $L''(t) = 0$. We conclude that $Y(h(t), t)$ and $(\nabla_X Y)(h(t), t)$ are both multiples of $X(h(t), t)$. Since $Y(u, t)$ is a Jacobi field as a function of u, it must equal $(au + b)X(u, t)$ for some constants a, b. Setting $u = 0, 1$ and applying π_*, it follows that $\gamma'(t) = b\alpha_t'(0) = (a+b)\alpha_t'(1)$; thus $a = 0$, α_t is a closed geodesic, and α_t is a reparametrization of γ. Clearly this makes α a reparametrization of γ and L constant.

10.2 REMARK. In the above proof, if $\delta : M_1 \to M_1$ is the deck transformation corresponding to α, then L is the displacement function of δ along $\gamma_1 : L(t) = d(\gamma_1(t), \delta\gamma_1(t))$. Thus we have shown that this displacement function g_δ is convex on M_1. This same result also can be obtained from Proposition 4.7 because $g_\delta(p)/\sqrt{2}$ is the distance in $M_1 \times M_1$ (with the product metric) from a point $(p, \delta(p))$ in the graph of δ to the diagonal submanifold $D = \{(p, p) \mid p \in M\}$. One needs to observe further that D and the graph of δ are totally geodesic submanifolds homometric to M_1.

If $t \to \alpha_t$ is the motion of a geodesic loop α along a geodesic ray γ, we define α to be γ-*monotone* provided $t \to |\alpha_t|$ is monotone nonincreasing. If $K < 0$, this length is strictly decreasing or, in the exceptional case, constant.

10.3 REMARK. With the notation as above, the following are equivalent:

(1) α is γ-monotone.
(2) γ_1 is g_δ-monotone.
(3) $\delta \circ \gamma_1$ is asymptotic to γ_1.

The equivalence of (2) and (3) follows from Corollary 9.3.

The set of elements of $\pi_1(M, m)$ containing a γ-monotone geodesic loop is

denoted $\pi_1(M, m, \gamma)$. These subsets of $\pi_1(M, m)$ are called *ray subgroups*, the terminology being justified by

10.4 PROPOSITION. *Let $m \in M$ (complete, $K \leq 0$).*
(1) *For each ray γ starting at m, $\pi_1(M, m, \gamma)$ is a subgroup of $\pi_1(M, m)$.*
(2) *$\pi_1(M, m)$ is the union of the ray subgroups $\pi_1(M, m, \gamma)$ for all such rays γ.*

Proof. (1) Suppose $\alpha, \beta \in \pi_1(M, m, \gamma)$ and that δ, ε are the corresponding deck transformations. Then for each $p \in M_1$

$$
\begin{aligned}
g_\delta(p) + g_\varepsilon(p) &= d(p, \delta p) + d(p, \varepsilon p) \\
&= d(p, \delta p) + d(\delta p, \delta \varepsilon p) \\
&\geq d(p, \delta \varepsilon p) = g_{\delta\varepsilon}(p).
\end{aligned}
$$

Since g_δ and g_ε are bounded on γ_1, so is $g_{\delta\varepsilon}$; that is, $\alpha\beta$ is γ-monotone.

Since a loop and its reverse have the same length, α^{-1} is γ-monotone.

(2) We must show that any geodesic loop at m is monotone on some ray, or equivalently, that for any deck transformation δ, some ray γ_1 from m is g_δ-monotone. If g_δ has no minimum, this follows from Proposition 9.7. If g_δ (and hence also g_δ^2) has a minimum, then by Proposition 4.2, δ translates a geodesic τ_1, so g_δ is constant on τ_1. By Proposition 9.8, the asymptote to τ_1 from m_1 is also g_δ-monotone.

We now check that there is no essential change in the ray groups when the base point m of $\pi_1(M, m)$ is changed. In fact, we can distinguish the ray subgroups in the deck transformation group $\pi_1(M)$, which is a basepoint-free version of $\pi_1(M, m)$. Indeed, it is clear from Remark 10.3 and Proposition 9.8 that one of these subgroups of $\pi_1(M)$ is associated to each asymptote class A of rays in M_1. Thus we define the *asymptote class subgroup* $\pi_1(M, A)$ to be those deck transformations δ such that A is stable under δ. If we choose a base point $m_1 \in M_1$, then there is an associated isomorphism between $\pi_1(M)$ and $\pi_1(M, \pi(m_1))$, under which the subgroup $\pi_1(M, [\gamma_1])$ corresponds to the ray subgroup $\pi_1(M, \pi(m_1), \pi \circ \gamma_1)$. Here γ_1 is any ray starting at m_1 and $[\gamma_1]$ is its asymptote class. In the light of this discussion the following is a restatement of Proposition 10.4.

10.5 PROPOSITION. *Let M be complete with $K \leq 0$. Then for every deck transformation δ of the simply connected covering M_1 of M, there is an asymptote class of rays in M_1 which is stable under δ. Every ray subgroup of $\pi_1(M, m)$ corresponds, by choice of base point in M_1 above m, to the stability subgroup of some asymptote class of rays in M_1.*

10.6 COROLLARY. *The inner automorphisms of $\pi_1(M, m)$ permute the ray subgroups.*

Let $\pi_1(M, m, \gamma)_0 = \{[\alpha] \in \pi_1(M, m, \gamma) \mid \lim |\alpha_t| = 0 \text{ on } \gamma\}$. The argument in Proposition 10.4 (1) shows that $\pi_1(M, m, \gamma)_0$ is also a subgroup. This group can be

nontrivial, for example, when γ runs to infinity on a tube whose diameter shrinks toward zero.

Now we assume that M has strictly negative curvature. We need the following lemma.

10.7 LEMMA. *If M has $K \leq c < 0$, then the distance between two asymptotic rays is zero.*

Proof. We may assume that M is simply connected, since a covering space projection is distance nonincreasing.

Let σ and τ be asymptotic rays in M, with τ having unit speed. Let β_t be the minimizing geodesic segment from $\tau(t)$ to σ, parametrized on I. Define a rectangle $r: I \times [0, \infty) \to M$ by $r(u, t) = \beta_t(u)$. For sufficiently large t, say $t \geq 0$, r is C^∞. As in Lemma 10.1, we obtain the same formula for LL'', where $L(t) = |\beta_t|$. This time $Y(\cdot, t)$ is not only a Jacobi field along each β_t, but also satisfies the end condition at the totally geodesic submanifold $\sigma(0, \infty)$. Thus $\langle \nabla_Y Y, X \rangle(0, t) = 0$ and $\| Y \|(\cdot, t)$ is an increasing function (of u) on I. Applying Rolle's theorem again we obtain inequalities

$$LL''(t) \geq -K(X, Y) \| X \wedge Y \|^2 (h(t), t)$$
$$\geq -cL^2(t) \| Y(0, t) \|^2.$$

Let $f(v) = \| Y(0, v) \|$. Since $Y(0, \cdot)$ is the velocity field of the parametrization $r(0, \cdot)$ of σ, the triangle inequality gives

$$\int_0^t f(v) \, dv = d(r(0, 0), r(0, t))$$
$$\geq -d(r(0, 0), \tau(0)) + d(\tau(0), \tau(t)) - d(\tau(t), r(0, t))$$
$$\geq t - 2L(0).$$

However, by Schwarz's inequality (with second function $g = 1$), for $t > 2L(0)$,

$$t \int_0^t f(v)^2 \, dv \geq \left(\int_0^t f(v) \, dv \right)^2 \geq (t - 2L(0))^2.$$

Hence we also have, for some $k > 0$ and $t > 2L(0)$,

$$\int_0^t f(v)^2 \, dv \geq t - k.$$

Now noting that $f(v) \leq \| Y(1, v) \| = 1$, we have

$$\frac{1}{c} (L'(t) - L'(0)) + \int_0^t L(v) \, dv = \int_0^t \left(L(v) + \frac{1}{c} L''(v) \right) dv$$

$$\leq \int_0^t L(v)(1 - f(v)^2) \, dv$$

$$\leq L(0) \int_0^t (1 - f(v)^2) \, dv$$

$$\leq kL(0).$$

Since $\lim L'(t) = 0$, this shows that the integral $\int_0^\infty L(v)\,dv$ converges, and hence $\lim L(t) = 0$.

10.8 COROLLARY. *If M (complete, $K \leq c < 0$) is simply connected and $\phi: M \to M$ is an isometry, then for every geodesic ray γ, $\lim_{s \to \infty} g_\phi(\gamma(s))$ is an invariant of the asymptote class of γ, where $g_\phi(m) = d(m, \phi(m))$. Moreover, this limit has only two possible values, ∞ and $\inf g_\phi$. If g_ϕ has no minimum, there is a unique g_ϕ-monotone asymptote class. If g_ϕ has a nonzero minimum, then ϕ translates a geodesic τ and the only g_ϕ-monotone asymptote classes are $[\tau^+]$ and $[\tau^-]$.*

Proof. If g_ϕ goes to ∞ on γ, the same obtains for every ray asymptotic to γ. Thus we may assume $\inf g_\phi \circ \gamma < \infty$. Suppose τ is asymptotic to γ. Let $\varepsilon > 0$. Then by the triangle inequality, for every s, t such that $d(\tau(t), \gamma(s)) < \varepsilon$

$$|d(\gamma(s), \phi\gamma(s)) - d(\tau(t), \phi\tau(t))| \leq d(\gamma(s), \tau(t)) + d(\phi\tau(t), \phi\gamma(s)) < 2\varepsilon.$$

Hence $\inf g_\phi \circ \tau = \inf g_\phi \circ \gamma$.

Now let σ be a geodesic ray on which $a = \inf g_\phi \circ \sigma = \inf g_\phi < \infty$ and suppose that on some other ray γ we have $\inf g_\phi \circ \gamma = b < \infty$. By what we have just proved, we may assume that σ and γ intersect, so by Lemma 9.10 there is a geodesic τ such that τ^+ is asymptotic to σ and τ^- is asymptotic to γ. Thus g_ϕ is bounded, hence constant on τ, and $b = a$.

The final remarks follow from Propositions 4.2 and 9.11.

10.9 PROPOSITION. *If δ is a deck transformation of the simply connected covering $\pi: M_1 \to M$ (complete, $K \leq c < 0$), then either (1) $\inf g_\delta = 0$ or (2) δ translates a geodesic.*

Proof. Suppose $\inf g_\delta = a > 0$. Let γ be a g_δ-monotone geodesic, parametrized with unit speed so that $d(\delta\gamma(0), \gamma) < a$. We also may assume that the foot $\gamma(t_1)$ of $\delta\gamma(0)$ on γ is in the positive half, that is, $t_1 > 0$, since otherwise we replace δ by δ^{-1} and γ by $\delta\gamma$. Define a sequence $\{t_i\}$ recursively by:

$$\gamma(t_i) \text{ is the foot of } \delta\gamma(t_{i-1}) \text{ on } \gamma.$$

Let $d_i = d(\delta\gamma(t_i), \gamma(t_{i+1})) = d(\delta\gamma(t_i), \gamma)$ and $a_i = d(\gamma(t_i), \delta\gamma(t_i))$, so $a_i \geq a$. Since $\delta\gamma$ is drawing steadily closer to γ, it follows by continuity that $t_{i+1} > t_i$, and hence that a_i and d_i are nonincreasing. From the right triangle with vertices $\gamma(t_i)$, $\delta\gamma(t_i)$, and $\gamma(t_{i+1})$ we have $a - d_i \leq a_i - d_i \leq t_{i+1} - t_i \leq a_i$. Thus $\lim t_i = \infty$, and hence $\lim a_i = a$, $\lim d_i = 0$, and $\lim (t_{i+1} - t_i) = a$.

We now show that the series $\sum d_i$ converges, and hence that the sequence $\{\pi\gamma(t_i)\}$ converges to some point $m \in M$. As in Lemma 10.7, with $\tau = \delta\gamma$, $\sigma = \gamma$, we have that the infinite integral of $L(t) = d(\delta\gamma(t), \gamma)$ converges and L is decreasing, so

$$\int_0^\infty L(t)\,dt \geq \sum_{i=1}^\infty L(t_{i+1})(t_{i+1} - t_i)$$

$$\geq (a - d_1) \sum L(t_{i+1})$$

$$= (a - d_1) \sum d_{i+1}.$$

It follows that the sequence $\{\delta^{-i}\gamma(t_i)\}$ converges in M_1 to some point m_1 above m. But g_δ is invariant under δ, so $g_\delta(m_1) = \lim g_\delta(\delta^{-i}\gamma(t_i)) = \lim g_\delta(\gamma(t_i)) = a$. Thus g_δ has a minimum point m_1, so δ translates a geodesic by Proposition 4.2.

10.10 THEOREM. *Let α be a nonconstant geodesic loop at $m \in M$ (complete, $K \leqq c < 0$).*
(1) *If α is not freely homotopic to a closed geodesic then*:
 a. *α is monotone on exactly one ray γ from m,*
 b. *$\pi_1(M, m, \gamma)_0 = \pi_1(M, m, \gamma)$,*
 c. *γ diverges.*
(2) *If α is freely homotopic to a closed geodesic τ, then*:
 a. *α is monotone on exactly two rays ρ and σ from m,*
 b. *$\pi_1(M, m, \rho) = \pi_1(M, m, \sigma)$, and this group is infinite cyclic,*
 c. *τ is a limit cycle of ρ and σ with ρ asymptotic to τ, σ to its reverse curve τ^{-1},*
 d. *$\pi_1(M, m, \rho)_0 = \pi_1(M, m, \sigma)_0 = \{1\}$.*

Together with Proposition 10.4, this gives

10.11 COROLLARY. *If M is a complete manifold with strictly negative curvature, then the fundamental group $\pi_1(M, m)$ is the disjoint (except for $\{1\}$) union of its ray groups.*

Proof. Let δ be the deck transformation corresponding to α. Since the geodesic loops in the free homotopy class of α are the projections of segments from p to δp, the cases correspond to the cases (1) and (2) of Proposition 10.9.

In (1), $\inf g_\delta = 0$, so by Corollary 10.8 there is a unique g_δ-monotone asymptote class. Thus $[\alpha]$ is in a unique ray group $\pi_1(M, m, \gamma)$ and $\lim |\alpha_t| = 0$, that is, $[\alpha] \in \pi_1(M, m, \gamma)_0$. To prove c, let B be any compact subset of M. There is an $\varepsilon > 0$ such that ε is a normal radius at each point of B. If $\gamma(t) \in B$, then we must have $|\alpha_t| > \varepsilon$. Thus for some t_0, $\gamma(t) \notin B$ for all $t > t_0$. To prove b, let β be a geodesic loop, $[\beta] \in \pi_1(M, m, \gamma)$. If $[\beta]$ is not in $\pi_1(M, m, \gamma)_0$, then by Proposition 10.9, the deck transformation of β would translate a geodesic τ. Hence γ would be asymptotic to the closed geodesic $\pi\tau$, contradicting c.

In (2), parts a and c are immediate consequences of Corollary 10.8. Moreover, d follows easily from b, so it remains to prove b. We first note that the subgroup of $\pi_1(M)$ which leaves invariant a given geodesic is infinite cyclic and is generated by a unique smallest positive translation. Thus we want to show that every member of $\pi_1(M, m, \rho)$ corresponds to a deck transformation which translates the same geodesic τ_1 which δ translates.

Suppose that $\phi \in \pi_1(M, A)$, $\phi \neq 1$, where $A = [\tau_1^+]$. By (1) b, since $\inf g_\delta \neq 0$, we also have $\inf g_\phi \neq 0$, and hence ϕ translates a geodesic γ_1 asymptotic to τ_1. We may suppose that both δ and ϕ are the smallest positive translators of τ_1 and γ_1. Then for every t, $\lim_{n \to \infty} d(\delta^n \tau_1(t), \gamma_1) = 0$, which shows that $\pi\tau_1(t)$ is in the closed

geodesic $\pi\gamma_1$. Thus by reparametrizing, $\pi\tau_1 = \pi\gamma_1$, and there is a deck transformation μ such that $\mu\tau_1 = \gamma_1$. But $\mu\delta\mu^{-1}$ translates the geodesic $\mu\tau_1$, in fact by a minimum positive amount, hence $\mu\delta\mu^{-1} = \phi$, and δ and ϕ translate by the same amount a. Now for $\varepsilon > 0$, let t be large enough so that $d(\tau_1(t), \gamma_1) = d(\tau_1(t), \gamma_1(s)) < \varepsilon/2$ for some s. Since $d(\tau_1(t+u), \gamma_1(s+u))$ is a convex function of u, bounded for u positive, we also have $d(\tau_1(t+a), \gamma_1(s+a)) < \varepsilon/2$. By the triangle inequality and the fact that ϕ is an isometry we have

$$d(\tau_1(t), \phi^{-1}\delta\tau_1(t)) < d(\gamma_1(s), \phi^{-1}\tau_1(t+a)) + \varepsilon/2$$
$$= d(\gamma_1(s+a), \tau_1(t+a)) + \varepsilon/2$$
$$< \varepsilon.$$

This shows that $\inf g_{\phi^{-1}\delta} = 0$. But by (1)b, the only deck transformation in $\pi_1(M, A)$ having this property is 1. Thus $\delta = \phi$ and $\tau_1 = \gamma_1$.

10.12 Proposition. *Let M be complete, $K \leq c < 0$; and let $G = \pi_1(M, m, \gamma)$ be a ray subgroup of $\pi_1(M, m)$. If $\beta \notin G$, then $\beta G \beta^{-1} \cap G = \{1\}$.*

Proof. Let δ be the deck transformation corresponding to β and let γ_1 be the lift of γ to the base point m_1 above m. By hypothesis $\delta \notin \pi_1(M, [\gamma_1]) = G_1$, so $\delta\gamma_1$ is not asymptotic to γ_1. If G_1 is of the second type, consisting of translations of a geodesic τ_1, then $\delta\tau_1^-$ is not asymptotic to τ_1^- either. But $\delta G_1 \delta^{-1}$ is the asymptote class subgroup $\pi_1(M, [\delta\gamma_1])$ (which also equals $\pi_1(M, [\delta\tau_1^-])$ in the second case), so it meets G_1 only in $\{1\}$ by the disjointedness properties of ray subgroups.

Geometrically this property of ray groups means that if a nonconstant geodesic loop α shortens as it moves out a ray γ, but β does not, then the geodesic loop in $[\beta\alpha\beta^{-1}]$ also does not shorten. The example of a riemannian product, with each factor having negative curvature and nontrivial homotopy, shows that this property fails if we assume merely $K \leq 0$, even with the additional assumption of negative mean curvature.

10.13 Corollary. *If M has strictly negative curvature, there are 0, 1, or ∞ nontrivial ray subgroups. Furthermore, if there are ∞, then there are either 0 or ∞ of each of the types (1) and (2) of Theorem 10.10.*

Proof. If there were a finite number ($\neq 0$, 1) of ray subgroups, then inner automorphism by any element would permute them, and hence have a power leaving them fixed. The same power of the element must be the identity, but only the identity has finite order. Furthermore, the types (1) and (2) are preserved under inner automorphisms (since $g_\delta \circ \phi = g_\mu$, where $\mu = \phi^{-1}\delta\phi$), so the same argument may be used with each type separately.

10.14 Remarks. In case there are ∞ ray subgroups of type (2), Corollary 10.13 does not necessarily mean that there are infinitely many distinct closed geodesics (as point sets). In fact, each closed geodesic is counted infinitely many times, because conjugate group elements are freely homotopic to the same closed geodesic

and each element has infinitely many conjugates. The corresponding condition when there are infinitely many type (1) subgroups, the existence of finitely many rays from a point such that every type (1) ray is asymptotic to one of them, definitely does occur; for example, a thin tube of strictly negative curvature can be attached to a compact surface; or two such tubes can be attached to the hyperbolic plane. However, we suspect that the existence of two closed geodesics or a closed geodesic and a type (1) ray implies the existence of infinitely many closed geodesics.

Some other simple consequences of Proposition 10.12 are: (1) No proper ray subgroup is normal.

(2) If two elements α, β of $\pi_1(M, m)$ commute, they are in the same ray subgroup.

(3) Every commutative subgroup of $\pi_1(M, m)$ is contained in a single ray subgroup.

(4) If $\pi_1(M, m)$ has nontrivial center, then $\pi_1(M, m)$ itself is a ray group, that is, there is only one nontrivial ray subgroup.

We shall say that a riemannian manifold M is *full* provided each nontrivial free homotopy class of loops in M contains a closed geodesic. Equivalently, M is full if every deck transformation of its simply connected covering space translates a geodesic. Compact manifolds are full, but there are, of course, many full non-compact manifolds. Preissmann [8] has proved for complete manifolds with $K < 0$: If the deck transformation group consists of translations, then every commutative subgroup is infinite cyclic. We shall elaborate this result below in Proposition 10.16. By Theorem 10.10, M (complete, $K \le c < 0$) is full if and only if $\pi_1(M, m, \gamma)_0 = \{1\}$ for every geodesic ray γ from m which diverges.

10.15 PROPOSITION. *If M (complete, $K \le 0$) contains a compact totally convex set A, then M is full.*

Proof. Let $\alpha \in \pi_1(M, m)$ and let δ be the deck transformation of M_1 corresponding to α under some choice of base point $m_1 \in M_1$. Then the geodesic loops in the same free homotopy class as α are just the projections of the segments from p to $\delta(p)$, as p runs through M_1. Choose a sequence $\{p_i\}$ such that $\lim g_\delta(p_i) = \inf g_\delta$. Let $\rho_1 : M \to A_1 = \pi^{-1}(A)$ be the continuous retraction of Proposition 3.4. Since ρ_1 is distance-nonincreasing and commutes with δ, $\{q_i = \rho_1(p_i)\}$ is a sequence in A_1 having $\lim g_\delta(q_i) = \inf g_\delta$ also. Let τ_i be the projection of the segment from q_i to δq_i. Since $\tau_i(0) \in A$, which is compact, we may assume that the initial velocities $\tau_i'(0)$ converge, and hence that the loops τ_i converge to a loop τ. The convergence of geodesic segments is uniform, so τ is in the same free homotopy class as α, and clearly $|\tau| = \inf g_\delta$. Thus there is a lift of τ to some segment from, say, p_0 to δp_0, and since p_0 is a minimum point of g_δ, τ is a closed geodesic, by Proposition 4.2.

10.16 PROPOSITION. *If M is a complete full manifold with $K < 0$, then the fundamental group $\pi_1(M)$ has the properties*:

A. *It is a disjoint union of infinite cyclic subgroups.*

B. *Two commuting elements belong to the same cyclic subgroup.*

C. *Each of the subgroups in A is an asymptote class subgroup of type* (2) *as in Theorem* 10.10.

Proof. (By the method of Preissmann.) For a given geodesic τ in M_1 that is translated by some deck transformation, the members of $\pi_1(M)$ which translate τ form an infinite cyclic subgroup D_τ. Since M is full each deck transformation δ is in some D_τ. Since $K < 0$, Proposition 4.2 shows that each $\delta \neq 1$ is in at most one D_τ. This establishes A and C.

To prove B, suppose $\delta \in D_\tau$ and that μ commutes with $\delta \neq 1$. Then $\delta \mu \tau = \mu \delta \tau = \mu \tau$, so $\mu \tau$ is the geodesic translated by δ, that is, $\mu \tau = \tau$. Hence $\mu \in D_\tau$.

10.17 REMARKS. (1) If τ is a translated geodesic in M_1, then every tubular neighborhood of τ having uniform radius is a region of strictly negative curvature. Indeed, it is covered by isometric translates of a compact region. Thus we could have used the method of Theorem 10.10. This also shows that the ray of a nontrivial ray subgroup has the corresponding closed geodesic as a limit cycle.

(2) Property B is a restatement of Proposition 10.12 for this situation, and in the presence of A is equivalent to the assertion that if $x^m y = y x^n$ then x and y are in the same cyclic subgroup.

(3) In an abstract group G the *height* of $x \in G$ is $h(x) = \sup \{n \mid \text{there is } y \in G \text{ such that } y^n = x\}$. It requires rather ordinary group-theory technique to prove that property A for G is equivalent to the combination of the following three properties:

a. G is torsion free.

b. Every element of G has finite height.

c. If $x^m = y^n$ (m, n not both 0), then x and y lie in the same cyclic subgroup.

(4) If a nontrivial group G satisfies A, then the number of component cyclic subgroups is either 1 or ∞. (In the presence of B this follows as in the proof of Corollary 10.13, but it is true in general.)

(5) The fundamental group of the Klein bottle F is isomorphic to the group with two generators x, y and one relation $x^2 = y^2$. This description shows that it does not satisfy A. Since F admits a flat metric, we have that the fundamental groups of the following types of manifolds do not necessarily satisfy A:

a. Compact (in particular, full), $K \leq 0$. *Example. F with the flat metric.*

b. Not full, $K \leq c < 0$. *Example. $R \times_{e^t} F$.*

REFERENCES
(Also see §2.)

1. R. L. Bishop and R. J. Crittenden, *Geometry of manifolds*, Academic Press, New York, 1964.

2. H. Busemann, *The geometry of geodesics*, Academic Press, New York, 1955.

3. T. Frankel, *On theorems of Hurwitz and Bochner*, J. Math. Mech. **15** (1966), 373–377.

4. R. Hermann, *Homogeneous riemannian manifolds of nonpositive sectional curvature*, Nederl. Akad. Wetensch. Proc. Ser. A66 **25** (1963), 47–56.

5. S. Kobayashi, *Fixed points of isometries*, Nagoya Math. J. **13** (1958), 63–68.

6. S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, Interscience, New York, 1963.

7. B. O'Neill, *The fundamental equations of a submersion*, Michigan Math. J. **13** (1966), 459–469.

8. A. Preissmann, *Quelques propriétés globales des espaces de Riemann*, Comment. Math. Helv. **15** (1942–1943), 175–216.

University of Illinois,
    Urbana, Illinois
University of California,
    Los Angeles, California