

AN IMPROVED TOPONOGOV COMPARISON THEOREM FOR NONNEGATIVELY CURVED MANIFOLDS

DOUG ELERATH

INTRODUCTION

The main result of this paper is a global comparison theorem of Toponogov type which gives improved estimates in open nonnegatively curved manifolds. Previously, estimates on such a manifold could only be made by comparison with \mathbf{R}^2 . It is apparent that further improvements can be made, which lead towards an integral formulation of the theorem.

The application presented in section 5 is intended primarily as an example of the usefulness of this comparison theorem, and not as the optimal approach to the conjecture of Cheeger and Gromoll. It does in fact seem to me that this approach must be stretched to its limit if it is intended to be used in an effort to prove the conjecture.

The following is a list of the contents of this paper:

1. PRELIMINARIES

- 1.1. Notation
- 1.2. Curvature and conjugate points
- 1.3. Rauch-Berger comparison theorem

2. THE CUT LOCUS ON A FLATTENING SURFACE OF REVOLUTION

- 2.1. Preliminaries
- 2.2. Results
- 2.3. Example

3. HINGES, TRIANGLES AND BRANCHED COVERINGS

- 3.1. Definitions and notation
- 3.2. Hinges in flattening surfaces
- 3.3. Hinges in branched coverings of flattening surfaces

Communicated by H. Samelson, July 21, 1978. This represents part of a Ph.D. dissertation written under the guidance of Professor Jeff Cheeger at Stony Brook. It is a pleasure for the author to acknowledge the influence that Professor Cheeger has had on this work.

4. GLOBAL COMPARISON THEOREMS

4.1. Definitions and notation

4.2. Two global comparison theorems

5. MANIFOLDS DIFFEOMORPHIC TO EUCLIDEAN SPACE

5.1. The soul of a manifold

5.2. Compact half spaces and shriveled souls.

Portions, particularly §§3 and 4, are rather technical; but this is perhaps unavoidable.

1. PRELIMINARIES

This section recalls certain elementary aspects of differential geometry which are cited in the body of this paper, and establishes notation to be used throughout. A proof will be given only for Corollary 1.3.2 which, to the best of the author's knowledge, is not standard. All proofs and definitions which are omitted from this section are most easily found in Cheeger and Ebin [2].

1.1. Notation

M^n will denote a complete n -dimensional Riemannian manifold with metric \langle , \rangle and Levi-Civita connection ∇ . Unless otherwise specified, all geodesics in M will be assumed to be normal; $\|\gamma'\| = 1$, where $\gamma'(t)$ denotes the tangent vector to γ at $\gamma(t)$. If x and y are points in M , then $\text{Cur}(x, y)$, $\text{Geo}(x, y)$ and $\text{Geo}_m(x, y)$ will denote respectively those curves, geodesics and globally minimal geodesics which begin at x and end at y . Elements of $\text{Cur}(x, y)$ are required to be at least piecewise smooth. Thus if $\gamma \in \text{Cur}(x, y)$ is parameterized on $[0, b]$, then $\gamma(0) = x$ and $\gamma(b) = y$. Similarly, if $N \subseteq M$, then $\text{Cur}(N)$, $\text{Geo}(N)$ and $\text{Geo}_m(N)$ will denote respectively those curves, geodesics and globally minimal geodesics which have images contained in N . Again, elements of $\text{Cur}(N)$ are required to be at least piecewise smooth. If the parameter range of any curve γ is not specified, it is understood to be $[0, L[\gamma]]$, where $L[\gamma]$ denotes the length of γ . Finally, if $\gamma \in \text{Geo}(x, y)$ and $\eta \in \text{Geo}(x, z)$, we let $\angle(\gamma, \eta)(x) = \arccos\langle \gamma'(0), \eta'(0) \rangle$, $0 \leq \angle(\gamma, \eta) \leq \pi$.

If $\gamma : [0, b] \rightarrow M$, define $-\gamma : [0, b] \rightarrow M$ by $-\gamma(t) = \gamma(b - t)$. If $\gamma_1 \in \text{Cur}(x, y)$ and $\gamma_2 \in \text{Cur}(y, z)$ are parameterized on $[0, b_1]$ and $[0, b_2]$ respectively, define $\gamma_1 \vee \gamma_2 : [0, b_1 + b_2] \rightarrow M$ by

$$\gamma_1 \vee \gamma_2 = \begin{cases} \gamma_1(t) & \text{if } t \in [0, b_1], \\ \gamma_2(t - b_1) & \text{if } t \in [b_1, b_1 + b_2]. \end{cases}$$

Thus $\gamma_1 \vee \gamma_2 \in \text{Cur}(x, y)$.

Sometimes we shall, without further warning, let the same symbol denote both a curve and its image.

If $N \subset M$ is any subset and $r \in \mathbf{R}$ is any positive real number, $T_r(N) = \{x \in M \mid d(x, N) < r\}$ is an (open) tubular neighborhood of N . Here, of course, d denotes distance in the metric space structure induced by the connection.

1.2. Curvature and conjugate points

If $x \in M$, and σ is a plane in M_x , then $K(\sigma)$ will denote the sectional curvature of M at x determined by any two vectors spanning σ . The notation $K(x)$ will occur only in conjunction with an inequality sign, and will denote either $\inf\{K(\sigma)\}$ or $\sup\{K(\sigma)\}$ over all planes σ in M_x according as we have $K(x) \geq$ or $K(x) \leq$.

If $\gamma \in \text{Geo}(x, y)$, we will say that γ is *free of conjugate points* when we mean that $\gamma(t)$ is not conjugate to x along γ for any t in the domain of γ .

1.2.1. Lemma. *Let M^n and M_0^{n+k} be Riemannian manifolds, let $\gamma : [0, b] \rightarrow M$ and $\gamma_0 : [0, b] \rightarrow M_0$ be normal geodesics, and suppose that $K(\gamma_0(t)) \geq K(\gamma(t))$ for all $t \in [0, b]$. Then if γ_0 is free of conjugate points, so is γ .*

1.2.2. Remark. If M and M_0 are both 2-dimensional this lemma follows immediately from the Sturm comparison theorem for second order ordinary differential equations. See, for example, [10, p. 118]. Thus, despite the application of this relatively modern result [ca. 1950] to surfaces in §2, the techniques there should be considered entirely classical.

1.3. Rauch-Berger comparison theorem

Suppose that x and y are points in M^n . If $\gamma \in \text{Geo}(x, y)$, we will say that γ is *free of focal points* when we mean that $\gamma(t)$ is not a focal point of the $(n - 1)$ -dimensional embedded submanifold defined by restricting \exp to a sufficiently small neighborhood of $0 \in \gamma'(0)^\perp \subset M_x$.

1.3.1. Theorem (Berger). *Let M^n and M_0^{n+k} be Riemannian manifolds. Let $\gamma : [0, b] \rightarrow M$ and $\gamma_0 : [0, b] \rightarrow M_0$ be normal geodesics with γ_0 free of focal points. Assume for each $t \in [0, b]$, each $v \in M_{\gamma(t)}$ and each $v_0 \in M_{0\gamma_0(t)}$ that the sectional curvatures of the sections σ and σ_0 spanned respectively by $(\gamma'(t), v)$ and $(\gamma_0'(t), v_0)$ satisfy $K(\sigma_0) \geq K(\sigma)$. Let $T(t) = \gamma'(t)$ and $T_0(t) = \gamma_0'(t)$, and let J and J_0 be Jacobi fields along γ and γ_0 respectively satisfying*

- (1) $(\nabla_T J)(0)$ and $(\nabla_{T_0} J_0)(0)$ are tangent to γ and γ_0 respectively,
- (2) $\|\nabla_T J\|(0) = \|\nabla_{T_0} J_0\|(0)$,
- (3) $\langle T, J \rangle(0) = \langle T_0, J_0 \rangle(0)$, and
- (4) $\|J\|(0) = \|J_0\|(0)$.

Then $\|J\|(t) \geq \|J_0\|(t)$ for each $t \in [0, b]$.

The following corollary, although well known, appears nowhere in the literature, so a proof of it is given here.

1.3.2. Corollary. *Let $\gamma : [0, b] \rightarrow M^n$ and $\gamma_0 : [0, b] \rightarrow M_0^{n+k}$ be normal geodesics, and let E and E_0 be parallel unit vector fields along γ and γ_0 respectively with $\langle E, \gamma' \rangle = \langle E_0, \gamma_0' \rangle$. Suppose that $c : [0, b] \rightarrow M$ and $c_0 : [0, b] \rightarrow M_0$ are defined by*

$$c(t) = \exp_{\gamma(t)}(f(t)E(t)),$$

$$c_0(t) = \exp_{\gamma_0(t)}(f(t)E_0(t)),$$

where $f : [0, b] \rightarrow \mathbf{R}$ is smooth. Let $\eta_t : [0, 1] \rightarrow M$ and $\eta_{0t} : [0, 1] \rightarrow M_0$ be defined by

$$\eta_t(s) = \exp_{\gamma(t)}(sf(t)E(t)),$$

$$\eta_{0t}(s) = \exp_{\gamma_0(t)}(sf(t)E_0(t)).$$

Assume that for each $(t, s) \in [0, b] \times [0, 1]$, $K(\eta_t(s)) \leq K(\eta_{0t}(s))$, and that for each $t \in [0, b]$, η_{0t} is free of focal points. Then $L[c] \geq L[c_0]$.

Proof. Since c and c_0 are both parameterized on $[0, b]$, it suffices to compare the lengths of their tangent vectors.

If $t_1 \in [0, b]$ is fixed, and $\alpha : (-\epsilon, \epsilon) \times [0, 1] \rightarrow M$ is a variation given by $\gamma(t, s) + \eta_{t_1+t}(s)$, then $V(s) = (\partial\alpha/\partial t)(0, s)$ is a Jacobi field along η_{t_1} with end values $V(0) = \gamma'(t_1)$ and $V(1) = c'(t_1)$. Note further that since $\nabla_{\gamma'}E = 0$ and $\eta_t'(0) = f(t)E(t)$,

$$\begin{aligned} (\nabla_{\eta_t} V)(0) &= [\nabla_{V(0)} f(t)E(t)](t_1) \\ &= [V(f(t))E(t) + f(t)\nabla_{\gamma'}E(t)](t_1) \\ &= f'(t_1)E(t_1). \end{aligned}$$

Similarly, we find that $(\nabla_{\eta_{0t}} V_0)(0) = f'(t_1)E_0(t_1)$. This shows that V and V_0 satisfy condition (1) of Theorem 1.3.1. Evidently they satisfy the other conditions, and thus $\|V\|(1) \geq \|V_0\|(1)$. Hence $\|c'(t_1)\| \geq \|c_0'(t_1)\|$ for any $t_1 \in [0, b]$, and thus $L[c] \geq L[c_0]$.

Remark. Note that $\langle E, \gamma' \rangle$ and $\langle E_0, \gamma_0' \rangle$ were *not* required to be zero in this corollary.

2. THE CUT LOCUS ON A FLATTENING SURFACE OF REVOLUTION

Interest in the cut locus on surfaces dates back to Poincaré [15], and Myers [13], [14]. They both proved that the cut locus of a point x on a surface S can contain no closed curve, and that the end points of the cut locus of x are

conjugate to x . In addition, Myers proved that on an analytic surface the cut locus of any point is a tree with finitely many nodes in any compact subset of S .

To illustrate the notion of the cut locus, Myers describes the cut locus on several standard surfaces such as the plane, the sphere and the ellipsoid, and then states that "examples of well-known simply connected surfaces on which the (cut) locus assumes a complicated, but determinable, form are naturally hard to give." It is in fact extremely difficult to calculate the cut locus on a specific surface unless the surface is quite nice. Gluck and Singer [8] have recently shown that any smooth manifold of dimension ≥ 2 can be given a Riemannian metric with a non-triangulable cut locus, and they construct such a metric on the 2-sphere.

Previously, "quite nice" has commonly meant a quadric surface, or perhaps a cylinder or torus. Even in the case of a surface of revolution, where geodesics can be explicitly exhibited in integral form, the cut locus is not generally known. This problem was, however, solved for paraboloids and hyperboloids in a beautiful paper by von Mangoldt [12]. In this paper he sidesteps the task of calculating the elliptic integrals defining geodesics on these surfaces, and computes the conjugate locus directly. The cut locus is then immediately apparent.

In this section we will extend the results of von Mangoldt to the class of flattening surfaces of revolution. The techniques will not be his; but, as mentioned in §1, the techniques are essentially classical.

2.1. Preliminaries

Let $(r, \theta, z) : \mathbf{R}^3 \rightarrow [0, \infty) \times [0, 2\pi) \times \mathbf{R}$ be cylindrical coordinates. Let $S^2 \subset \mathbf{R}^3$ be a smooth surface of revolution about the z -axis, with the induced metric, such that $(r, \theta)|_S$ is 1-1. Let $p = r|_S^{-1}(0)$ be the vertex of S , and let $\rho : S \rightarrow \mathbf{R}$ be defined by $\rho(s) = d(p, s)$, where the distance d is measured in the surface S .

Using (a) the integral expression for a geodesic on a surface of revolution and (b) Clairaut's theorem (see for example [7]), one can easily deduce the following qualitative description of the behavior of geodesics on S .

2.1.1. Description. On each infinite normal geodesic $\gamma : \mathbf{R} \rightarrow S$ there is a unique point, denoted σ_γ , which is nearest to the vertex of S , about which γ is symmetric, and at which γ is tangent to a parallel of S . The two branches of γ

proceed in either direction from this point and spiral in opposite senses around the axis of S heading monotonically towards infinity, so that $\rho \circ \gamma$ satisfies a maximum principle.

Ling [11] provides a more detailed discussion of geodesics on a slightly more restricted class of surfaces.

2.1.2. Definition. Such a surface of revolution S will be said to be *flattening* if $dK/d\rho \leq 0$, where K denotes the Gauss curvature of the surface and ρ was defined above.

Thus we add to our notion of a surface of revolution the requirement that the curvature decreases as distance from the vertex increases.

2.1.3. Notation. We let S be a flattening surface of revolution, and set

- (1) $S_+ = S \cap \theta^{-1}(0, \pi)$,
- (2) $P_x =$ the parallel ($z = \text{constant}$) through a point $x \in S$,
- (3) $\mu_x =$ the meridian ($\theta = \text{constant}$) through a point $x \in S$ with $\mu_x(0) = P$,

(4) $\hat{\mu}_x =$ the meridian opposite x , i.e., $\hat{\mu}_x(t) = \exp_p(-t\mu'_x(0))$. If $\gamma \in \text{Geo}(S)$, $\text{domain}(\gamma) = \mathbf{R}$, then

- (5) $\tau_\gamma = \inf r(\gamma(t))$, $t \in \mathbf{R}$,
- (6) $\sigma_\gamma =$ that unique point in S on γ for which $r(\sigma_\gamma) = \tau_\gamma$. If $\gamma \in \text{Geo}(S)$, $\text{domain}(\gamma) \neq \mathbf{R}$, then

(7) $\bar{\gamma} \in \text{Geo}(S)$ denotes the unique extension of γ so that $\text{Im}(\gamma) \subset \text{Im}(\bar{\gamma})$ and $\text{domain}(\bar{\gamma}) = \mathbf{R}$,

$$(5') \tau_\gamma = \tau_{\bar{\gamma}}$$

$$(6') \sigma_\gamma = \sigma_{\bar{\gamma}}$$

Suppose that $\gamma \in \text{Geo}(S_+)$, $\text{domain}(\gamma) = [0, b]$, γ not a meridian.

(8) $\theta_\gamma : [0, b] \rightarrow [0, \pi)$ by $\theta_\gamma(t) = |\theta(\gamma(t))|$. That θ_γ is 1-1 follows from Description 2.1.1.

(9) $F_\gamma : \text{Im}(\theta_\gamma) \rightarrow \rho(\gamma)$ by $F_\gamma = \rho \circ \gamma \circ \theta_\gamma^{-1}$, so that γ can be viewed as the graph of F_γ .

2.1.4. Definition. If $\gamma, \eta \in \text{Geo}(S_+)$, we say that η *lies below (above)* γ if (i) $\theta(\gamma) \subset \theta(\eta)$, and (ii) for each $t \in \text{domain}(\gamma)$, $\rho(\eta \cap \mu_{\gamma(t)}) \leq (\geq) \rho(\gamma(t))$. The same terminology will be used if γ is simply a point. Note that if η lies below γ , it is not necessarily the case that γ lies above η .

Finally, we need the following result which is implicit in the paper [14] of Myers.

2.1.5. Lemma (Myers). *Suppose M^2 is a smooth complete surface, and $x \in M$, $y \in M$. Given $g \in \text{Geo}_m(x, y)$, $h \in \text{Geo}_m(x, y)$ with $g \cup h$ the boundary of a simply connected region D , then there exists $z \in \text{Cl}(D)$ such that x and z are conjugate along a geodesic $g_0 \in \text{Geo}_m(\text{Cl}(D))$.*

2.2 Results

In Lemmas 2.2.2, 2.2.3 and 2.2.4 we assume the following data.

2.2.1. Data (see Fig. 1).

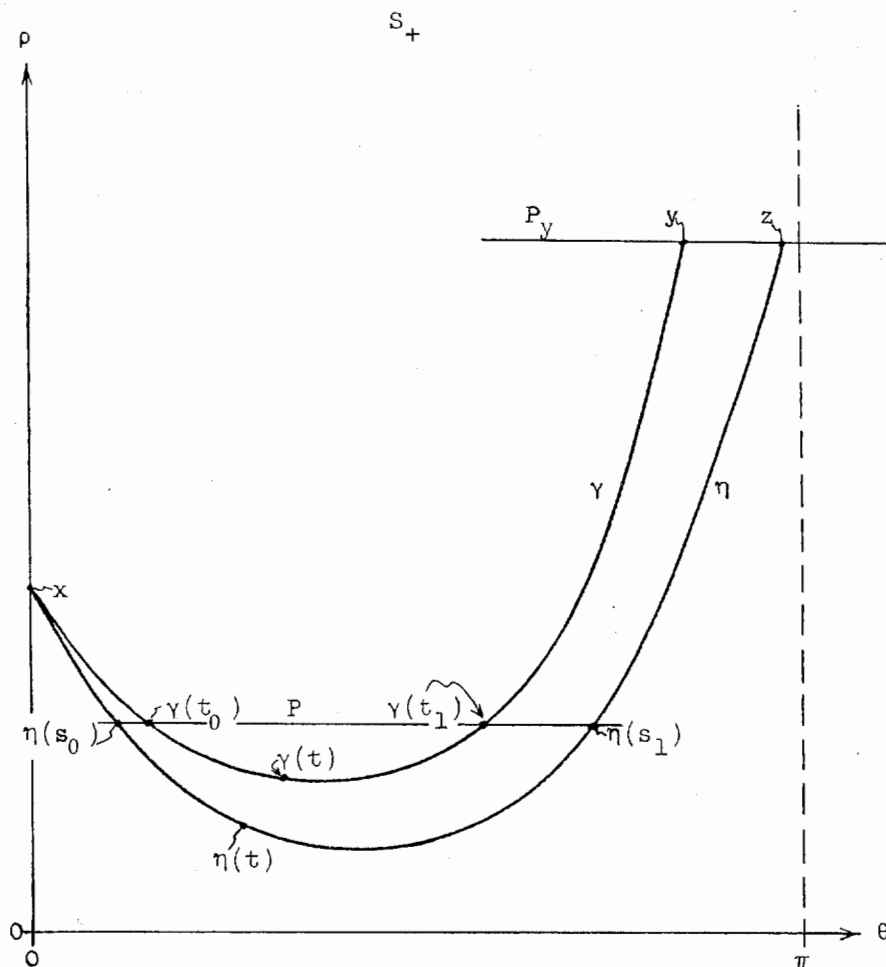


FIG. 1

- (1) $x \in \partial S_+$ with $\theta(x) = 0$;
- (2) $y \in S_+$ with $\rho(x) \leq \rho(y)$;
- (3) $\gamma \in \text{Geo}_m(x, y)$, $\gamma : [0, a] \rightarrow S$;
- (4) $z \in P_y \cap S_+$ is such that $\theta(z) > \theta(y)$ (S_+ is open);
- (5) $\eta \in \text{Geo}_m(x, z)$, $\eta : [0, b] \rightarrow S$.

The following three technical lemmas explore in detail the behavior of the geodesics γ and η .

2.2.2. Lemma. $\gamma, \eta \subset S_+$ and η lies below γ .

Proof. The first claim follows immediately upon observing that (i) S is symmetric with respect to reflection through the plane determined by ∂S_+ , and (ii) γ and η are minimal. Thus if, for example, h is any geodesic from x to z , we may reflect all portions of h not in S_+ through the plane determined by ∂S_+ . If $h \not\subset Cl(S_+)$, the result is a non-smooth curve from x to z of the same length as h , and hence h could not have been minimal. (This shows that S_+ is convex.)

Since $\theta(z) > \theta(y)$, $\gamma, \eta \subset S_+$ and $\theta_\gamma, \theta_\eta$ are 1-1, we have $\theta(\gamma) \subset \theta(\eta)$. Since both γ and η are minimal, $\eta \cap \gamma = \{x\}$. Thus η lies either above γ or below γ . Let $H(t) = F_\eta(t + \theta_\eta(0))$ and $G(t) = F_\gamma(t + \theta_\gamma(0))$ map $\theta(\gamma)$ to \mathbf{R} . If η lies above γ , $H > G$, and in particular $H(\theta(y)) > G(\theta(y))$. Then $H(0) < H(\theta(z)) < H(\theta(y))$, while $0 < \theta(y) < \theta(z)$. This is of course impossible since H cannot achieve a maximum at an interior point (cf. Description 2.1.1).

2.2.3. Lemma. $L[\gamma] < L[\eta]$.

Proof. Suppose that P_y is parameterized, not by arc length, but such that $\theta(z) - \theta(P_y(s)) = s$, and let $z_s = P_y(s)$. From $\eta'(\theta) > 0$, $P'_y(\theta) < 0$ and $P'_y(\rho) = 0$, it follows that $\langle \eta', P'_y(z) \rangle < 0$. Now choose a smooth variation $\alpha : (-\delta, \delta) \times [0, b] \rightarrow S$ of η with $\alpha(s, 0) = x$, $\alpha(s, b) = z_s$ and $\alpha(0, t) = \eta(t)$. Then $\langle \partial/\partial s(\alpha), \eta' \rangle|_{s=0; t=b} = \langle P'_y, \eta' \rangle(z) < 0$, and so the first variation formula implies that $\partial/\partial s|_{s=0} L[\alpha, (s, \cdot)] < 0$.

Since both Lemma 2.2.2 and the argument of the preceding paragraph apply as well to each $\eta_s \in \text{Geo}_m(x, z_s)$ for all relevant s , i.e., when $\theta(y) < z_s \leq \theta(z)$, it follows that $L[\eta_s]$ decreases monotonically as s increases. Thus $L[\gamma] < L[\eta]$.

2.2.4. Lemma. $\rho(\gamma(t)) > \rho(\eta(t))$ for each $t > 0$ in the domain of γ .

Proof. Let P be any parallel with $\tau_\gamma < \rho(P) \leq \rho(y)$. Let t_1 be such that $\gamma(t_1) \in P$ and $F'_\gamma(t_1) > 0$, and thus for all $t > t_1$, $\rho(\gamma(t)) > \rho(P)$. (That such a t_1 exists is clear from Data 2.2.1 and Description 2.1.1). Let s_1 be similarly defined in terms of η . The argument given in Lemma 2.2.3 applies, mutatis mutandis, to show that $t_1 < s_1$.

If $P \cap \gamma$ is but a single point, $P \cap \eta$ will likewise be a single point, and necessarily these intersections will be the $\gamma(t_1)$ and $\eta(s_1)$ described in the previous paragraph ($\rho(y)\rho(z) = \rho(x)$). Thus, since P lies above $\eta[[0, s_1]$, $\rho(\eta(t_1)) < \rho(\eta(s_1)) = \rho(\gamma(t_1))$.

If $P \cap \gamma$ is two points (the only other case), $P \cap \eta$ will also be two points, and it must be that $F'_\gamma(0) < 0$. Define t_0 so that $\gamma(t_0) \in P$ and $F'_\gamma(t_0) < 0$, and let s_0 be similarly defined in terms of η . Thus if $t_0 < t < t_1$, $\rho(\gamma(t)) < \rho(P)$; while if $s_0 < t < s_1$, $\rho(\eta(t)) < \rho(P)$. Let $w = \gamma(t_0)$ and let P_w be parameterized so that $\theta(w) - \theta(P_w(s)) = s$. Let $w_s = P_w(s)$ and let $\gamma_s \in$

$\text{Geo}_m(x, w_s)$ for $0 \leq s < \theta(w)$. It is easily seen that $F'_{\gamma_s}(\theta(w_s)) < 0$: Rotate γ so that $\gamma(t_0)$ passes through w_s ; it is then clear that γ_s is trapped above this rotated version of γ , and thus $F'_{\gamma_s}(\theta(w_s)) < F'_{\gamma}(t_0) < 0$. Thus $\langle \gamma'_s, P' \rangle_{w_s} < 0$, and using a first variation argument as in Lemma 2.2.3 we see that $s_0 < t_0$ (the geodesics are parameterized by arc length).

Now given any $t \in \text{domain}(\gamma)$ we construct $P_{\gamma(t)}$ with s_0, s_1, t_0 and t_1 as above. The foregoing argument implies that $s_0 < t_0 \leq t \leq t_1 < s_1$, and thus that $\rho(\eta(t)) < \rho(\eta(s_1)) = \rho(\gamma(t_1)) = \rho(\gamma(t))$. q.e.d.

In the following we let $C(x)$ denote the cut locus of x .

2.2.5. Theorem. *Let S be a flattening surface of revolution and $x \in S$. Then $C(x) \subset \hat{\mu}_x$.*

Proof. Without loss of generality we may suppose that $x \in \partial S_+, \theta(S_+) = (0, \pi)$, and $\theta(x) = 0$. If $C(x) \cap S_+ \neq \emptyset$, then the convexity of S_+ (shown in Lemma 2.2.2), the characterization of cut points given in Lemma 5.2 of [2], and Lemma 2.1.5 together imply that there exists a $y \in S_+$ such that x and y are conjugate along a minimal geodesic in S_+ . By relabeling, if necessary, we may suppose that $\rho(x) < \rho(y)$. We will now see that such a situation, i.e., x conjugate to such a y in S_+ , is not possible, and thus that $C(x) \cap S_+ = \emptyset$. The theorem is then apparent.

So, suppose we are given Data 2.2.1 on S . Then, since S is flattening, $\rho(\eta(t)) < \rho(\gamma(t))$ for $t \in \text{domain}(\gamma)$ implies that $K(\eta(t)) > K(\gamma(t))$. Moreover, since η is minimal, x can have no conjugate points along η for a distance equal to $L[\eta] > L[\gamma]$. Thus Lemma 1.2.1 shows that y is not conjugate to x along γ , and we are done.

2.3. Example

This example shows that Theorem 2.2.5 is sharp in the following sense: Given $\delta > 0$, there exist a surface of revolution S and a point $x \in \partial S_+$ such that (i) $dK/d\rho < 0$ except on a set $E \subset S$, (ii) $m(E) < \delta$, where m denotes Lebesgue measure, and (iii) $C(x) \cap S_+ \neq \emptyset$. Equivalently, by multiplying the metric by a constant, we can replace (ii) by (ii') $dK/d\rho|_E < \delta$.

We will construct a surface which is only piecewise smooth, and in fact does not satisfy our requirements for a surface of revolution (cf. §2.1.). But standard approximation theorems, see for example Aleksandrov and Zalgaller [1], imply that this is sufficient.

2.3.1. Let $f : [0, \infty) \rightarrow \mathbb{R}^2$ be the curve

$$f(t) = \begin{cases} (t, 0) & 0 \leq t < 1, \\ (1, t - 1) & 1 \leq t, \end{cases}$$

and let S be the surface generated by revolving f about the vertical axis. See Fig. 2. Let $\gamma : \mathbf{R} \rightarrow S$ be a geodesic with $0 < \tau_\gamma < 1$, and $\sigma_\gamma = (\tau_\gamma, 0, 0)$. Then for some $\varepsilon > 0$, $\gamma|[-t_0, t_0] \subset S_+$, where $t_0 = (1 - \tau_\gamma^2)^{1/2} + \varepsilon$, and both $\gamma(t_0)$ and $\gamma(-t_0)$ are in the cylindrical part of S . Thus there is another geodesic $\eta \in \text{Geo}(\gamma(-t_0), \gamma(t_0)) \cap \text{Geo}(S_+)$ which is distinct from γ , and so $C(\gamma(t_0)) \cap S_+ \neq \emptyset$. Now we can find an approximation to S for which (i) and (ii), or (i) and (ii') hold, and a point x on the approximation such that (iii) holds.

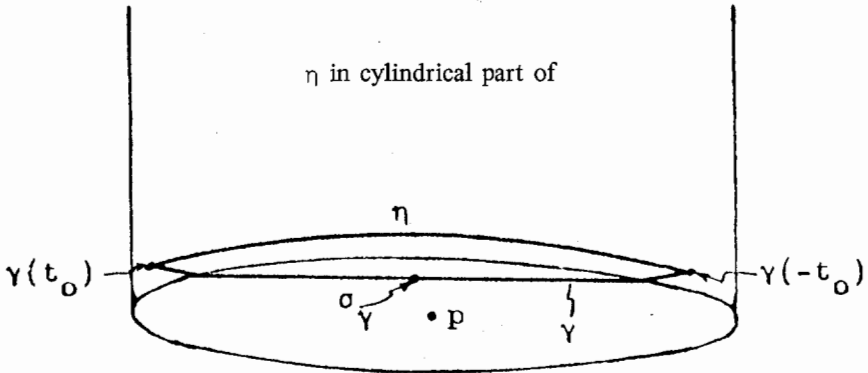


FIG. 2

2.3.2. With a little more care, (i), (ii) and (ii') can all be made to hold simultaneously. Here we let $f_\theta : [0, \infty) \rightarrow \mathbf{R}^2$ be the curve

$$f_\theta(t) = \begin{cases} (1 - \cos(t), \sin(t)) & 0 \leq t \leq \theta \leq \pi/2, \\ (1 - \cos(\theta), t - \theta + \sin(\theta)) & \theta \leq t, \end{cases}$$

and let S_θ be the surface generated by revolving f about the vertical axis. Just as in §2.3.1 we find, for a given $\theta < \pi/2$, that there is an $x \in \partial S_+$ with $C(x) \cap S_+ \neq \emptyset$.

If we now approximate S_θ by a smooth surface of revolution T_θ which satisfies (i) and (ii') (T_θ may differ from S_θ by a conformal factor), then, if θ is sufficiently near $\pi/2$, T_θ will satisfy (i), (ii), (ii') and (iii).

3. HINGES, TRIANGLES AND BRANCHED COVERINGS

In this section we define hinges and geodesic triangles in a Riemannian manifold, and then prove several technical lemmas which examine hinges more closely on flattening surfaces and their branched coverings. The idea is to replace part (1) of the proof of Toponogov's theorem as presented in [2] with a lemma appropriate to a flattening surface of revolution. The reasons for going to the branched coverings are somewhat subtle, and will be indicated later in the paper.

3.1. Definitions and notation

Throughout this section, any geodesic written γ_i is assumed to be parameterized on $[0, b_i]$, and we let $e_i = \gamma_i(b_i)$. Furthermore, if $\{\gamma_i^\alpha\}$ is a family of geodesics which is parameterized by α with each γ_i^α parameterized on $[0, b_i]$, then set $e_i^\alpha = \gamma_i^\alpha(b_i)$.

If (M, d) is any metric space with $A \subset M, B \subset M$, let $d_A(B) = \sup\{d(A, x) | x \in B\}$, and define the Hausdorff distance between A and B by

$$hd(A, B) = \max\{d_A(B), d_B(A)\}.$$

If S is a surface of revolution and $\gamma \in \text{Geo}(S)$, we may extend 2.1.3(8) and define θ_γ as follows: We may suppose that $0 \in \text{domain}(\gamma)$. If $x \in \text{range}(\gamma)$, we may orient S in \mathbb{R}^3 so that $x \in \partial S_+, \theta(x) = 0$. Let (a, b) be the domain of the connected component of $\gamma|_{S_+}$ such that $\gamma(a) = x$. Then we require that (i) $\theta_\gamma(0) = 0$ and (ii) if $t \in (a, b)$, $\theta_\gamma(t) - \theta_\gamma(a) = \theta(t)$. It is easily seen that θ_γ is 1-1 for any nonmeridional geodesic γ on S .

Let M be a Riemannian manifold.

3.1.1. Definition. A hinge in M is a triple $(\gamma_1, \gamma_2, \alpha)$ with

- (1) $\gamma_i \in \text{Geo}(M)$,
- (2) $e_0 = \gamma_1(0) = \gamma_2(0)$, and
- (3) $\sphericalangle(\gamma_1, \gamma_2)(e_0) = \alpha$.

By a family of hinges $\{(\gamma_1^\alpha, \gamma_2, \alpha)\}$ we mean that for each α in some interval $(\gamma_1^\alpha, \gamma_2, \alpha)$ is a hinge in M with each γ_1^α parameterized on $[0, b_1]$. Note that $e_1^\alpha = \gamma_1^\alpha(b_1)$ is a continuous function of α .

Sometimes, given a hinge $(\gamma_1, \gamma_2, \alpha_0)$, we will speak of increasing or decreasing α_0 by moving, say, γ_1 . This procedure may be viewed formally by constructing the family of hinges $\{(\gamma_1^\alpha, \gamma_2, \alpha)\}, \alpha \in [a, b]$, with $\alpha_0 \in [a, b]$. Then, if $\alpha_1 \in [a, \alpha_0)$, to "decrease α_0 to α_1 " means to consider $\gamma_1^{\alpha_1}$ to be the geodesic resulting from this "movement".

3.1.2. Definition. A triple of geodesics $(\gamma_0, \gamma_1, \gamma_2)$ on M is a *geodesic triangle (on M)* if $\gamma_i(b_i) = \gamma_{i+1}(0), i \in \mathbb{Z}_3$. Note that if each γ_i is minimal, then it is apparent that the γ_i satisfy the *triangle inequality*, i.e.,

$$L[\gamma_i] + L[\gamma_{i+1}] \geq L[\gamma_{i+2}], \quad i \in \mathbb{Z}_3.$$

3.2. Hinges in flattening surfaces

For the remainder of §3, S will denote a flattening surface of revolution with vertex p , R will be a fixed positive real number, and attention will be restricted to the compact ball $Cl(T_{2R}(p))$ in S .

The following lemma is a bit unwieldy, but not without purpose. In this lemma we define a number of constants $\delta_1, \dots, \delta_5$ associated to a given

flattening surface of revolution. Having five instead of just one allows us to more easily point out exactly what hypotheses are needed in later propositions.

3.2.1. Lemma. *Let S be a flattening surface of revolution and fix $R \in \mathbf{R}$.*

(1) *There exists $\delta_1 > 0$ so that if $\gamma \in \text{Geo}(T_{2R}(p))$ with $L[\gamma] < d(\gamma(0), p) + 2\delta_1$, then γ is free of conjugate points.*

(2) *There exists $\delta_2 > 0$ so that if $\gamma \in \text{Geo}(T_{2R}(p))$ with $L[\gamma] < \delta_2$, then γ is free of focal points.*

Let G denote the set of normal geodesics in $T_R(p)$ such that if $\gamma \in G$, then $L[\gamma] < d(\gamma(0), p) + \delta_1$. Each $\gamma \in G$ is parameterized on $[0, L[\gamma]]$.

(3) *There exists $\delta_3 > 0$ such that if $\gamma \in G$ and η is any geodesic with $\text{hd}(\gamma, \eta) > \delta_3$, then η is unique between its end points among all those geodesics whose hd distance from γ is less than δ_3 .*

(4) *There exists $\delta_4 > 0$ such that if $\gamma \in G$ and $z \in T_{\delta_4}(\gamma(0))$, then there exists $\eta \in \text{Geo}(\gamma(L[\gamma]), z)$ with $\text{hd}(\eta, \gamma) < \delta_3$ and $L[\eta] < L[\gamma] + \delta_1$. We may furthermore require that $\angle(\gamma, \eta)(\eta(0)) < \pi$ unless $\gamma \vee \eta \in \text{Geo}(\gamma(0), z)$ (this is used only in Lemma 3.2.3).*

(5) *Given $\epsilon > 0$ and $\varphi \in (0, \pi/2]$, there exists a $\delta_5 \in (0, \delta_4)$ such that the following holds: Let $(\gamma_1, \gamma_2, \alpha)$ be a hinge with $\gamma_1 \in G$, $L[\gamma_2] < \delta_5$ and $\alpha \in [\varphi, \pi - \varphi]$. Let η be the unique geodesic from e_1 to e_2 whose hd distance from γ_1 is less than δ_3 . Recall that e_i was defined in §3.1. Then, up to reparameterization, $\eta(t) = \exp_{\gamma_1(t)}(t)f(t)E(t)$, where E is the parallel unit vector field arrived at by parallel translating γ_2 along γ_1 , and $f: [0, b] \rightarrow [0, \epsilon)$ is smooth.*

Proof. In each case we shall produce the δ for an arbitrary geodesic. Then a uniform δ can be found using continuity and compactness arguments.

(1) Suppose $\gamma \in \text{Geo}(x, y)$, and let μ be the meridian segment from x to p . Since x and p are not conjugate along μ , μ can be extended a small amount beyond p and still be free of conjugate points ($C(x)$ is closed). Thus, assuming that μ has been so extended, $L[\mu] > d(x, p)$ and μ contains the meridian segment from x to p . If $p \in \gamma$, we are evidently done. Otherwise, since μ is minimal from x to p , $d(\mu(t), p) < d(\gamma(t), p)$ for $0 < t \leq d(x, p)$, and thus $d(\mu(t), p) < d(\gamma(t), p)$ holds for $0 < t < d(x, p) + 2\delta_1$ for some $\delta_1 > 0$. Since S is flattening $K(\mu(t)) > K(\gamma(t))$, and the proof is completed by applying Lemma 1.2.1.

(2) This is clearly true on any compact manifold.

(3)–(5) These all follow easily from the facts that γ is free of conjugate points, and that $\exp_x v$ is a continuous function of both $x \in S$ and $v \in S_x$.

Remark. Any δ_i appearing in the remainder of this paper will be assumed to have been chosen in accordance with the above lemma.

3.2.2. Lemma. *Let $\{(\gamma_1^\alpha, \gamma_2, \alpha)\}$, $\alpha \in [a, b]$, be a family of hinges in S_+ . Then $d/d\alpha(d(e_1^\alpha, e_2)) > 0$ for all $\alpha \in [a, b]$.*

Proof. This is a straightforward analogue to step (1) in the proof of Theorem 2.2 in Cheeger and Ebin [2]. Here it is important to observe that if $x \in Cl(S_+)$, $C(x) \cap S_+ = \emptyset$.

3.2.3. Lemma. *Let $\{(\gamma_1^\alpha, \gamma_2, \alpha)\}$, $\alpha \in [0, \pi]$, be a family of hinges in $T_R(p) \subset S$. Suppose that*

- (1) $L[\gamma_1^\alpha] < d(e_0, p) + \delta_1$, and
- (2) $L[\gamma_2] < \delta_4$.

Let γ^α denote the unique geodesic from e_2 to e_1^α whose hd distance from γ_1^α is less than δ_3 . Then $(d/d\alpha)L[\gamma^\alpha] > 0$ for all $\alpha \in [0, \pi]$.

Proof. Since γ_1^α is free of conjugate points for $\alpha \in [0, \pi]$, $A = \{e_1^\alpha\}$ is a smooth submanifold (with boundary). Clearly, by choice of δ_4 , γ^α is also free of conjugate points for each α . This, using the fact that $T_{\delta_3}(\gamma_1^\alpha)$ (in the hd metric on $\text{Geo}(S)$) is open, implies that $L[\gamma^\alpha]$ is a smooth function of α . It is therefore reasonable to compute $(d/d\alpha)L[\gamma^\alpha]$.

The variation vector field for the variation $\{\gamma^\alpha\}$ is zero at e_2 and tangent to A at e_1^α . Since $\gamma_1^\alpha(e_1^\alpha)$ is orthogonal to A and $\angle(\gamma_1^\alpha, \gamma^\alpha)(e_1^\alpha) < \pi$ for $\alpha \in (0, \pi)$ (by choice of δ_4), γ^α is never orthogonal to A for $\alpha \in (0, \pi)$. Thus, using the first variation formula, we see that $(d/d\alpha)L[\gamma^\alpha] \neq 0$. But it is quite clear that

$$L[\gamma^0] = |L[\gamma_1] - L[\gamma_2]| < L[\gamma_1] + L[\gamma_2] = L[\gamma^\pi],$$

so that $(d/d\alpha)L[\gamma^\alpha] > 0$.

3.3. Hinges in branched coverings of flattening surfaces

If S is a flattening surface of revolution with vertex p , S^* will denote the infinite-sheeted branched covering space of S , branched over p , with the induced geometry. For computational purposes we may view S^* in the following manner.

Suppose \mathbf{R}^2 is represented in polar coordinates (r, θ) , and let $H = [0, \infty) \times \mathbf{R}$. Define $\pi_0 : H \rightarrow \mathbf{R}^2$ by

$$\pi_0(x, y) = (x, y \bmod 2\pi),$$

and topologize H so that π_0 is continuous. Now consider the diagram

$$\begin{array}{ccc} S^* & \xrightarrow{(\rho, \theta)^*} & H \\ \pi \downarrow & & \downarrow \pi_0 \\ S & \xrightarrow{(\rho, \theta)} & \mathbf{R}^2 \end{array}$$

where S^* is the (ρ, θ) -pullback of H . Let the geometry of S^* be the π -pullback of the geometry on S , and $p^* = \pi^{-1}(p)$. It is clear that the topology on S^* which is induced by the geometry is the same as the topology which S^* inherits as a pullback of H . Note that S^* is not geodesically complete at p^* , but that otherwise geodesics are simply lifts of geodesics in S . Note also that minimal paths in S^* are either lifts of minimal geodesics in S , or else can be written as $\mu_1 \vee \mu_2$ where μ_1 and μ_2 are meridian segments in S^* , μ_1 ending and μ_2 beginning at p^* .

Let $\rho^* = (\rho, \theta)^*$ followed by projection into the first factor of H , and $\theta^* = (\rho, \theta)^*$ followed by projection into the second factor of H , and note that as a map from S^* to H , $(\rho^*, \theta^*) = (\rho, \theta)^*$. Finally, if $\gamma \in \text{Geo}(S^*)$, let $\theta_{\gamma^*} = \theta_{(\pi \circ \gamma)}$.

3.3.1. Lemma. On $S^* \setminus \{p^*\}$

- (1) (ρ^*, θ^*) is one to one, and hence a homeomorphism between $S^* \setminus \{p^*\}$ and $H \setminus (\{0\} \times \mathbf{R})$,
- (2) each simple closed path determines a bounded and an unbounded component in S^* , $p^* \notin$ bounded component, and
- (3) there are no closed geodesics.

Proof. (1) Since (ρ, θ) is one to one, so is (ρ^*, θ^*) . Since π and (ρ, θ) are open maps, so is (ρ^*, θ^*) .

(2) This is standard, using the homeomorphism of (1).

(3) Suppose $\gamma : [0, b] \rightarrow S^*$, $\gamma \in \text{Geo}(x, x)$. Then using (2) it is clear that $\theta_{\gamma^*}(0) = \theta_{\gamma^*}(b)$, which is impossible since $\theta_{\gamma^*} (= \theta_{(\pi \circ \gamma)})$ is one to one. q.e.d.

Let S_+^* denote some connected component of $\pi^{-1}(S_+)$. The following lemma along with Lemma 3.3.5 provide the primary motivation for working in S^* rather than in S . This will be discussed further in §4.

3.3.2. Lemma. If $x \in S_+^*$ and $y \in S_+^*$, then $\text{Geo}(x, y)$ contains precisely one element, and $\text{Geo}(x, y) \subset \text{Geo}(S_+^*)$.

Proof. That there is some $\gamma \in \text{Geo}(x, y)$, $\gamma \subset S_+^*$, is clear since the geometry on S^* is the π -pullback of the geometry on S , and the minimal geodesic between $\pi(x)$ and $\pi(y)$ remains in S_+ . That $\text{Geo}(x, y) \cap \text{Geo}(S_+^*)$ contains only one element follows immediately from Theorem 2.2.5 which describes the cut locus on S . If $\eta \in \text{Geo}(x, y)$, $\eta \neq \gamma$, then it must leave S_+^* . But Lemma 3.3.1(2) then implies that θ_η^* is not one to one. q.e.d.

Much of the technical difficulty which we will encounter throughout the remainder of this paper is due to the fact that S^* is not complete. This next lemma provides us with the degree of completeness needed however. Note that we continue to apply the notation of §2.1 to S^* whenever the meaning is clear.

3.3.3. Lemma. *If $x \in S^*$, $y \in S^*$, $\gamma \in \text{Geo}(x, y)$, and γ lies below $z \in S^*$, then $\text{Geo}(x, z) \neq \emptyset$. If γ is also a minimal curve, so that $L[\gamma] = d(x, y)$, then there exists a minimal geodesic from x to z .*

Proof. Suppose $z \neq y$ and γ is not contained in a meridian (in either case the lemma is obvious).

Let μ denote the meridian through z , and ν the meridian through x with $\nu(0) = p^*$. See Fig. 3. Let $\varphi = \sphericalangle(\gamma, \nu)(x)$. Define a family of geodesics $\{\gamma_\alpha\}$, $\alpha \in [0, \varphi]$, each geodesic of which begins at x , lies above γ in some neighborhood of x , and such that $\sphericalangle(\gamma, \gamma_\alpha)(x) = \alpha$. Since, for $\alpha \neq \varphi$, γ_α is not a meridian, each γ_α can be extended indefinitely and in particular, at least for small α , until γ_α crosses μ . Furthermore, it is clear that $\gamma_\alpha \cap \mu$ depends continuously on α , and thus $A = \{\gamma_\alpha \cap \mu | \alpha \in [0, \varphi]\}$ is a connected set.

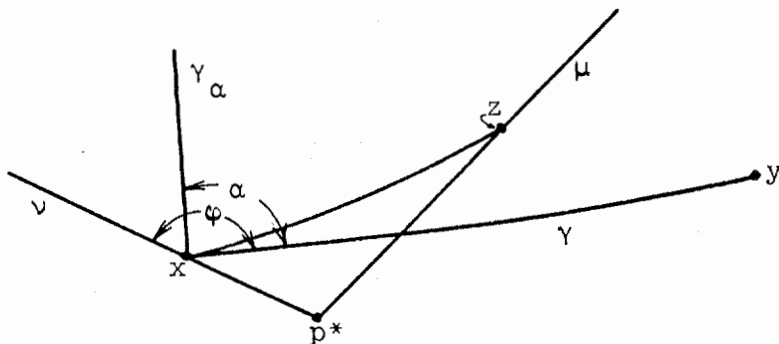


FIG. 3

It is clear that $\gamma_\varphi \cap \mu$ is empty and, by the continuous dependence of γ_α on α , that $C = \{\alpha | \gamma_\alpha \cap \mu = \emptyset\}$ is closed. Let a_0 be the smallest element in C , and let $\{\alpha_i\}$ be a sequence of real numbers, $0 < \alpha_i < \alpha_{i+1} < a_0$, which converge to a_0 . Let $m_i = \gamma_{\alpha_i} \cap \mu$, and we claim that $\rho^*(m_i) \rightarrow \infty$: if not, $m_i \rightarrow m \in \mu$; but again by continuous dependence of γ_α on α , we see that $\gamma_{a_0} \cap \mu = m$, a contradiction.

Thus, since $\rho^*(\gamma \cap \mu) < \rho^*(z)$ and A is connected, $\gamma_\alpha \cap \mu = z$ for some $\alpha \in [0, \varphi]$.

If γ is a minimal geodesic, let c be any minimal path from x to z . If c passes through p^* , then, since γ lies below z and meridians are the shortest paths from p^* , c will cross γ . Since $z \neq y$, either c or γ , or both, must continue to minimize beyond $c \cap \gamma$. But it is a standard fact that this cannot happen.

Thus the minimal path from x to z cannot pass through p^* , and so it must be a geodesic. That γ lies below this minimal geodesic is also quite easy to see. q.e.d

For $x \in S^*$ we call y a *cut point* of x if there is a geodesic from x to y which is minimal to, but not beyond, y . $C(x)$ will denote the collection of all cut points of x , and note that $C(x)$ may well be not-closed and/or disconnected. For example, let S be a semi-infinite cylinder capped with a hemisphere (Example 2.3.2 with $\theta = \pi/2$) take out the vertex, and let x lie in the interior of the hemisphere. Then $C(x) = \{p\} \cup \bar{\mu}_x|(a, \infty)$, where $a \neq 0$. $\bar{\mu}_x(a) \notin C(x)$ since x and $\bar{\mu}_x(a)$ are conjugate on S only along a geodesic through p . The same problem arises in S^* .

The following lemma is the natural extension of the corresponding fact on complete manifolds.

3.3.4. Lemma. *Let $\gamma \in \text{Geo}(S^*)$. Then $\gamma(t_0)$ is a cut point of $x = \gamma(0)$ if and only if one of the following holds for $t = t_0$, and none holds for any smaller value of t :*

- (1) $\gamma(t_0) = p^*$,
- (2) x and $\gamma(t_0)$ are conjugate along γ , or
- (3) there exists $\eta \in \text{Cur}(x, \gamma(t_0))$, $\eta \neq \gamma$, with $L[\eta] = L[\gamma]$.

Proof. Since, if γ is not a meridian nearby geodesics are also not meridians, the proof of the corresponding fact for complete manifolds, (see [2, Lemma 5.2]) applies. The only difference being that perhaps η passes through p^* , and is thus only necessarily piecewise geodesic.

3.3.5. Lemma. *Let $x \in S^*$ and $y \in C(x) \setminus \{p^*\}$. Then $d(x, y) \geq d(x, p^*) + 2\delta_1$, where δ_1 is chosen for the surface S with $S^* = \pi^*(S)$ and $R \gg d(x, y)$.*

Proof. Let $\gamma \in \text{Geo}_m(x, y)$. If x and y are conjugate along γ , then $\pi(x)$ and $\pi(y)$ are conjugate along $\pi(\gamma)$. Thus, by Lemma 3.2.1, $L[\pi(\gamma)] \geq d(\pi(x), p) + 2\delta_1$; and so $L[\gamma] \geq d(x, p^*) + 2\delta_1$.

Otherwise there exists $\eta \in \text{Cur}(x, y)$, $L[\eta] = L[\gamma]$. But then $\pi(\eta)$ and $\pi(\gamma)$ are distinct elements of $\text{Cur}(\pi(x), \pi(y))$ of the same length, and thus $\pi(\gamma)$ has quit minimizing prior to $\pi(y)$. Thus, by Lemma 5.6 of [2] and Lemma 3.2.1, we once again see that $L[\gamma] \geq d(x, p^*) + 2\delta_1$.

3.3.6. Remark. Note that Lemma 3.2.1 parts (1), (2) and (3), with obvious minor modifications, apply to S^* as well as to S . The incompleteness of S^* prevents (5) from being applied in S^* . Likewise Lemma 3.2.2 applies to hinges in some S_+^* , but Lemma 3.2.3 does not apply to S^* . Thus we need the following lemma on S^* .

3.3.7. Lemma. *Let $\{(\gamma_1^\alpha, \gamma_2, \alpha)\}$, $\alpha \in [a, b]$, be a family of hinges in S^* . Suppose*

- (1) $L[\gamma_1] \leq d(e_0, p^*)$,
- (2) $\gamma_2 \subset S_+^*$, S_+^* some given component of $\pi^{-1}(S_+)$, and

(3) for each $\alpha \in [a, b]$ and each $t \in [0, b_1]$ there exists a geodesic from $\gamma_1^\alpha(t)$ to e_2 without cut points.

Then $(d/d\alpha)(d(e_1^\alpha, e_2)) > 0$ for $\alpha \in (a, b)$.

Proof. Let $\gamma^{\alpha,t}$ denote the unique minimal geodesic from $\gamma_1^\alpha(t)$ to e_2 whose existence is assumed. From supposition (1) and Lemma 3.3.5 it follows that $A_t = \{\gamma_1^\alpha(t)\}$, $\alpha \in [a, b]$, is a smooth submanifold for each $t \in [0, b_1]$. Since γ_2 is in S_+^* and hence unique between its end points (Lemma 3.3.2), $(\gamma^{\alpha,t})'(0)$ is not orthogonal to A_t , lest $\gamma_1^\alpha|[0, t] \vee \gamma^{\alpha,t}$ be another geodesic from e_0 to e_2 . Furthermore, since e_2 has no cut points in $\{A_t\}$, $t \in [0, b_1]$, except perhaps p^* , it follows that $L[\gamma^{\alpha,t}]$ is a smooth function of both α and t . Thus, applying the first variation formula as in Lemma 3.2.3, $(\partial/\partial\alpha)L[\gamma^{\alpha,t}] \neq 0$.

It is easy to see that $\text{sgn}((\partial/\partial\alpha)L[\gamma^{\alpha,t}])$ is independent of t , and thus it suffices to prove the lemma for t small. But if t is sufficiently small the entire family $\{\gamma_1^\alpha\}$, $\alpha \in [0, \pi]$, is contained in S_+^* , in which case the lemma is apparent (cf. Lemma 3.2.2).

3.3.8. Remarks. (1) The geodesics required in Lemma 3.3.7(3) will always exist if $\gamma_1^\alpha \in S_+^*$ or γ_1^α lies below e_2 (Lemmas 3.3.2 and 3.3.3). The uniqueness is insured if $\{\gamma_1^\alpha\} \subset S_+^* \cup T_{d(e_2, p^*) + \delta_1}(e_2)$ (Lemmas 3.3.2 and 3.3.5).

(2) Note that in Lemma 3.3.7, γ_2 was allowed to be fairly long while in Lemma 3.2.3, γ_2 was required to be very short. Note however that in Lemma 3.2.3 we need not be as restrictive as in Lemma 3.3.7 to insure the existence of the γ^α .

4. GLOBAL COMPARISON THEOREMS

Toponogov's theorem ([18]; see [2] for proof in English) is a beautiful and powerful global generalization of the Rauch comparison theorem. It gives distance estimates on a Riemannian manifold M by comparison with a surface of constant curvature. Specifically,

Theorem (Toponogov). *Let M be a complete manifold with $K(x) \geq H \in \mathbf{R}$ for each $x \in M$. Let $(\gamma_1, \gamma_2, \alpha)$ be a hinge in M with γ_1 minimal and, if $H > 0$, $L[\gamma_2] \leq \pi/\sqrt{H}$. Let $(\bar{\gamma}_1, \bar{\gamma}_2, \alpha)$ be a hinge in the simply connected surface of constant curvature H with $L[\bar{\gamma}_1] = L[\gamma_1]$, and $\bar{\gamma}_i \in \text{Geo}(\bar{e}_0, \bar{e}_i)$. Then $d(e_1, e_2) \leq d(\bar{e}_1, \bar{e}_2)$.*

This theorem has provided estimates adequate for many important applications. It does, however, appear to be unnecessarily restrictive in the case of an open nonnegatively curved manifold. Since on any such manifold the curvature must come arbitrarily near zero (Bonnet's theorem), the comparison surface must be flat or negatively curved. If, for example, M is a paraboloid

and the hinge $(\gamma_1, \gamma_2, \alpha)$ has its vertex e_0 at the vertex of M , it is obvious that distance estimates on \mathbf{R}^2 will not be very accurate.

A natural question is: Can we improve the estimates on such a manifold? In the following we will show that we can, and that in fact, instead of a surface of constant curvature, an appropriate flattening surface of revolution may be used for the comparison surface. We actually prove two very similar such generalizations, the first serving as a lemma for the second.

4.1. Definitions and notation

Notational conventions established here will be used throughout this section without further comment.

4.1.1. Let M^n denote a complete, open, nonnegatively curved manifold, let $\delta \geq 0$ and $p \in M$. Let \bar{M} denote a flattening surface of revolution with vertex \bar{p} . We say that M and \bar{M} are δ -correspondent at p if, whenever $d(\bar{p}, \bar{x}) + \delta \geq d(p, x)$, then $K(\bar{x}) \leq K(x)$, where $x \in M$ and $\bar{x} \in \bar{M}$. That is, the curvature in \bar{M} falls off with respect to distance from \bar{p} faster than the curvature in M falls off with respect to distance from p . We will simply say that M and \bar{M} are correspondent at p when we mean that $\delta = 0$.

4.1.2. If M and \bar{M} are δ -correspondent at p , and $(\gamma_1, \gamma_2, \alpha)$ is a hinge in M with $e_0 = p$ or $e_1 = p$, then the hinge $(\bar{\gamma}_1, \bar{\gamma}_2, \alpha)$ in \bar{M} given by specifying that

- (1) $L[\bar{\gamma}_i] = L[\gamma_i]$, and
- (2) $e_i = p \Rightarrow \bar{e}_i = \bar{p}$, where $\bar{\gamma}_1 \in \text{Geo}(\bar{e}_0, \bar{e}_1)$ and $\bar{\gamma}_2 \in \text{Geo}(\bar{e}_0, \bar{e}_2)$.

is said to *correspond* to $(\gamma_1, \gamma_2, \alpha)$. Since $\bar{\gamma}_1$ is a segment of a meridian in \bar{M} which includes as one end point the vertex \bar{p} of \bar{M} , this correspondence is uniquely determined up to rotation and reflection of \bar{M} .

4.1.3. Given a hinge $(\gamma_1, \gamma_2, \alpha)$, we call a hinge (g_1, g_2, α) a *subhinge* if $g_2 \subset \gamma_2$. We say that the subhinge *faces inward (outward)* if $d(e_2, e_0) < (>)d(\epsilon_2, \epsilon_0)$, where $g_i \in \text{Geo}(\epsilon_0, \epsilon_i)$. See Fig. 4.

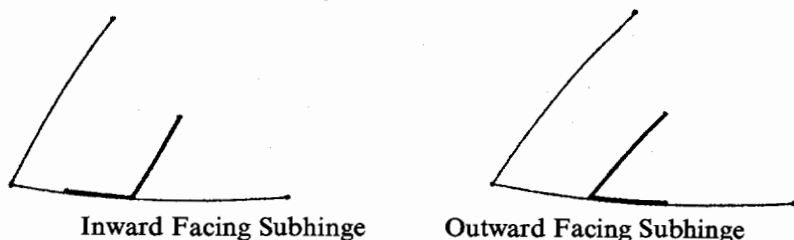


FIG. 4

For the remainder of this section, a *hinge* $(\gamma_1, \gamma_2, \alpha)$ in M will be assumed to have either $e_0 = p$ or $e_1 = p$ and γ_1 minimal. A *subhinge* will always be a

subhinge of such a hinge, with p not necessarily an endpoint of one of the two geodesics making up the subhinge.

Let $h = (\gamma_1, \gamma_2, \alpha)$ and $\bar{h} = (\bar{\gamma}_1, \bar{\gamma}_2, \alpha)$ be corresponding hinges in M and \bar{M} respectively. We say that the subhinges $(\gamma_1, \gamma_2, \alpha)$ of h and $(\bar{g}_1, \bar{g}_2, \alpha)$ of \bar{h} are *corresponding subhinges* if:

- (1) $L[g_i] = L[\bar{g}_i]$,
- (2) both face in the same direction, and
- (3) $d(e_0, \epsilon_0) = d(\bar{e}_0, \bar{\epsilon}_0)$, where ϵ_0 is as above.

4.1.4. Let $h = (\gamma_1, \gamma_2, \alpha)$ be a hinge in M and let $\Delta = (g_1, g_2, g_3)$ be a geodesic triangle in M with $g_2 \subset \gamma_2$. If $(\bar{\gamma}_1, \bar{\gamma}_2, \alpha)$ is a hinge in \bar{M} which corresponds to h , we say that a geodesic triangle $(\bar{g}_1, \bar{g}_2, \bar{g}_3)$ in \bar{M} *corresponds* to Δ if:

- (1) $L[\bar{g}_i] = L[g_i]$,
- (2) $\bar{g}_2 \subset \bar{\gamma}_2$, and
- (3) $d(e_0, g_2(t)) = d(\bar{e}_0, \bar{g}_2(t))$ as measured along γ_2 and $\bar{\gamma}_2$ respectively.

Finally, note that all of the above notions still make sense if we are working in \bar{M}^* instead of \bar{M} . The only situation which demands any care occurs when $(\gamma_1, \gamma_2, \alpha)$ is a hinge in \bar{M}^* with $e_0 = p^*$. Then one must be sure that γ_1 and γ_2 lie on the same branch of \bar{M}^* .

In the following lemma we assume that M and \bar{M} are δ -correspondent at $p \in M$, and that $(\gamma_1, \gamma_2, \alpha)$ and $(\bar{\gamma}_1, \bar{\gamma}_2, \alpha)$ are hinges (or subhinges) in M and \bar{M} respectively with $L[\gamma_i] = L[\bar{\gamma}_i]$. In this instance we do not assume that $e_i = p$ for $i = 0$ or 1 .

4.1.5. Lemma. *Suppose that $(\gamma_1, \gamma_2, \alpha)$ and $(\bar{\gamma}_1, \bar{\gamma}_2, \alpha)$ are as above with $0 < \theta_0 < \alpha < \pi - \theta_0$. Choose δ_5 with $\varphi = \theta_0$ and $\epsilon < \min\{\delta/2, \delta_2\}$, and suppose that $L[\gamma_2] < \delta_5$. Further suppose that $d(\gamma_1(t), p) \leq d(\bar{\gamma}_1(t), p)$, and let $\bar{\gamma}$ denote the unique geodesic from \bar{e}_2 to \bar{e}_1 whose hd distance from $\bar{\gamma}_1$ is less than δ_3 . Then $L[\bar{\gamma}] \geq d(e_2, e_1)$.*

Proof. It is quite straightforward to check that the hypotheses of Corollary 1.3.2 are satisfied.

Remarks. (1) Notice that if $e_1 = p$ and $d(\bar{e}_0, \bar{p}) \geq d(e_0, p)$, then the hypotheses of the lemma are satisfied. (2) This of course works as well in \bar{M}^* as long as the necessary geodesics exist. Rather than check this, we occasionally make the measurement in M and then pull back up to \bar{M}^* .

4.2. Two global comparison theorems

In both Theorem 4.2.1 and Theorem 4.2.2, M^n will denote a complete open nonnegatively curved n -dimensional Riemannian manifold, and \bar{M} a flattening surface of revolution such that M and \bar{M} correspond at $p \in M$.

4.2.1. Theorem Let $(\gamma_1, \gamma_2, \alpha_0)$ and $(\bar{\gamma}_1^*, \bar{\gamma}_2^*, \alpha_0)$ be corresponding hinges in M and \bar{M}^* respectively with γ_1 minimal, $e_1 = p$ and $\bar{\gamma}_2^* \subset \bar{M}_+^*$, assume that $\bar{\gamma}_i^* \in \text{Geo}(\bar{e}_0^*, \bar{e}_i^*)$. Let $\gamma_2 \in \text{Geo}_m(e_2, e_1)$. Then:

(A) $d(e_1, e_2) \leq d(\bar{e}_1^*, \bar{e}_2^*)$, and

(B) there exists a triangle $(c_1, \bar{\gamma}_2^*, c_3)$ in \bar{M}^* which corresponds to $(-\gamma_1, \gamma_2, \gamma_3)$, with

$$\angle(-c_1, \bar{\gamma}_2^*)(\bar{e}_0^*) \leq \angle(\gamma_1, \gamma_2)(e_0),$$

$$\angle(\bar{\gamma}_2^*, c_3)(\bar{e}_2^*) \leq \angle(\gamma_2, \gamma_3)(e_2).$$

4.2.2. Theorem. Let $(\gamma_1, \gamma_2, \alpha_0)$ and $(\bar{\gamma}_1, \bar{\gamma}_2, \alpha_0)$ be corresponding hinges in M and \bar{M} respectively with γ_1 minimal and $e_0 = p$. Let $\gamma_3 \in \text{Geo}_m(e_2, e_1)$. Then:

(A) $d(e_1, e_2) \leq d(\bar{e}_1, \bar{e}_2)$, and

(B) there exists a unique triangle $(c_1, \bar{\gamma}_2, c_3)$ in \bar{M} which corresponds to $(-\gamma_1, \gamma_2, \gamma_3)$, with

$$\angle(-c_1, \bar{\gamma}_2)(\bar{e}_0) \leq \angle(\gamma_1, \gamma_2)(e_0),$$

$$\angle(\bar{\gamma}_2, c_3)(\bar{e}_2) \leq \angle(\gamma_2, \gamma_3)(e_2).$$

Proof. The proofs of these two theorems are quite similar, and will be given simultaneously. The proof is divided into a number of steps each of which is further broken down into discussions of the difficulties particular to either Theorem 4.2.1 or Theorem 4.2.2.

Fix $R \in \mathbf{R}$, $R > 2(L[\gamma_1] + L[\gamma_2])$.

(1) We will make several simplifying assumptions, and then later show that the theorems as stated are correct.

Assume that for some $\delta > 0$, M and \bar{M} are δ -correspondent. If $x \in \gamma_2$ and $\gamma_x \in \text{Geo}_m(x, p)$, then assume that $\langle \gamma'_x, \gamma'_2 \rangle(x) \in (-1, 1)$. Note that, among other things, we have assumed that $\alpha_0 \in (0, \pi)$.

(2) Notation. As usual, notation established either in \bar{M} or \bar{M}^* will actually apply to both by insertion or deletion of a star (*).

Let $A = \{\text{arc cos}(\langle \gamma'_2, \gamma'_x \rangle(x)) \mid x \in \gamma_2 \text{ and } \gamma_x \in \text{Geo}_m(p, x)\}$. Let $\theta_s = \sup A$, and $\theta_i = \inf A$, as x and γ_x vary over all possibilities. Let $\theta_m = \min(\theta_i, \pi - \theta_s)$, and note that $\theta_m \in (0, \pi/2]$ by the assumptions of step (1) and the compactness of γ_2 . Now choose δ_5 with $\epsilon < \min\{\delta/2, \delta_2\}$ and $\varphi = \theta_m$.

Let $\{x_0 = e_0, x_1, \dots, x_{n-1}, x_n = e_2\}$ partition γ_2 , $x_i = \gamma(t_i)$, such that $t_i < t_{i+1}$ and $d(x_i, x_{i+1}) < \delta_5$. Choose $\sigma_i \in \text{Geo}_m(x_i, e_1)$ with $\sigma_0 = \gamma_1$, and assume that σ_i is parameterized on $[0, \lambda_i]$. Let $\{\bar{x}_i^*\}$ be a corresponding partition of $\bar{\gamma}_2^*$, with $\bar{x}_i^* = \bar{\gamma}_2^*(t_i)$. For $i < j$ define $\tau_{i,j} = \gamma_2[[t_i, t_j]$, and let $\bar{\tau}_{i,j}^*$ denote the corresponding pieces of $\bar{\gamma}_2^*$.

Call a geodesic σ in \bar{M}^* , with $\sigma(0) = \bar{x}_i^*$, interior if for all t sufficiently small the unique geodesic $\gamma^* \in \text{Geo}_m(e_0^*, \sigma(t))$ has $\angle(\gamma_1^*, \gamma^*)(e_0^*) < \alpha_0$. See Fig. 5. Let $\{\bar{\sigma}_i^*\}$ be a family of geodesics in \bar{M}^* so that (i) $\bar{\sigma}_i^*(0) = \bar{x}_i^*$, (ii) the subhinge $(\tau_{0,i}, \sigma_i, \beta_i)$ of $(\gamma_1, \gamma_2, \alpha_0)$ and the subhinge $(\bar{\tau}_{0,i}^*, \bar{\sigma}_i^*, \beta_i)$ of $(\bar{\gamma}_1^*, \bar{\gamma}_2^*, \alpha_0)$ are corresponding subhinges (note that this implicitly defines β_i), and (iii) so that each $\bar{\sigma}_i^*$ is an interior geodesic. Let $\alpha_i := \pi - \beta_i$.

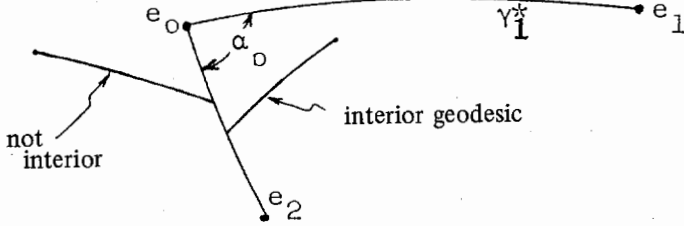


FIG. 5

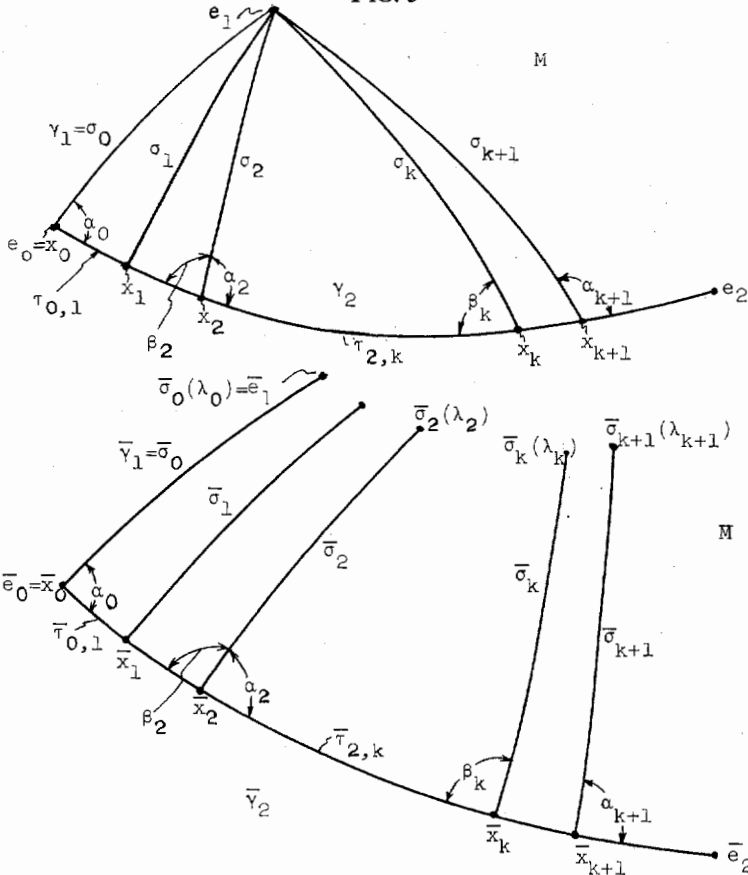


FIG. 6

Since $\angle(\bar{h}_1, \tau_{0,1})(x_0) \leq \alpha_0$ by construction, the only thing left to check is that $\angle(\bar{g}_1, -\bar{\tau}_{0,1})(\bar{x}_1) \leq \beta_1$. But it is quite easy to see that if β_1 is decreased by moving $\bar{\sigma}_1$ until $d(\bar{\sigma}_1(\lambda_1), \bar{x}_0)$, as measured along the nearby geodesic, is $L[\bar{\sigma}_0]$, then the resulting geodesic must be \bar{g}_1 . Since \bar{p} is not in the interior of the bounded region determined by $(-\bar{h}_1, \bar{\tau}_{0,1}, \bar{g}_1)$, the triangle, in the case of Theorem 4.2.1, can be lifted to \bar{M}^* .

Hence (B) is true for the triangle $(-\sigma_0, \tau_{0,1}, \sigma_1)$ in both Theorem 4.2.1 and Theorem 4.2.2.

Remark. In this step of Theorem 4.2.1 we had to work on \bar{M} since there was no *a priori* guarantee that \bar{h}_0 could be lifted to \bar{M}^* .

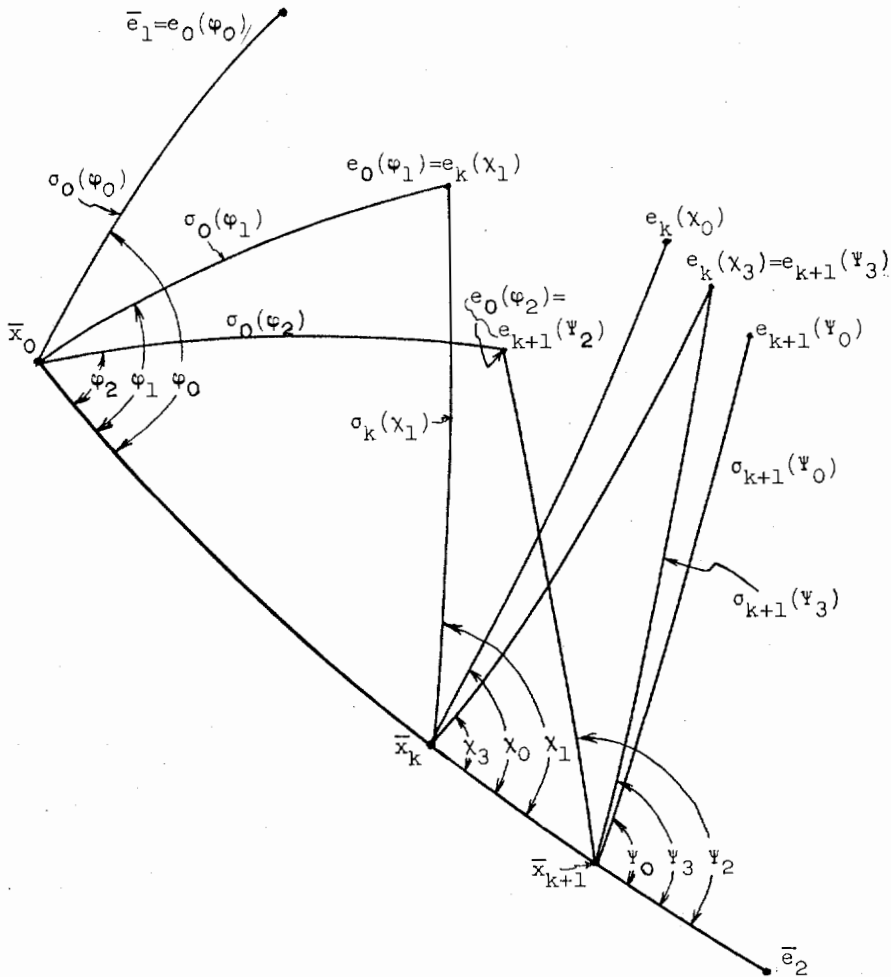


FIG. 8

(6) Now suppose that $d(e_0, x_k) \leq d(\bar{e}_0, \bar{x}_k)$ and that (B) is true for the triangle $(-\sigma_0, \tau_{0,k}, \sigma_k)$. In the following two parts several families of hinges shall be used. To avoid a surfeit of subscripted superscripts, we introduce the following notation:

$$\begin{aligned}\varphi_0 &= \alpha_0, \quad \chi_0 = \alpha_k, \quad \Psi_0 = \alpha_{k+1}; \\ \sigma_0(\varphi) &= \bar{\sigma}_0^{(\varphi_0 - \varphi)}, \quad \sigma_k(\chi) = \bar{\sigma}_k^{(\chi_0 - \chi)}, \\ c_{k+1}(\Psi) &= \bar{\sigma}_{k+1}^{(\Psi_0 - \Psi)} \quad (\text{cf. §3.1}); \\ e_0(\varphi) &= \sigma_0(\varphi)(\lambda_0), \quad e_k(\chi) = \sigma_k(\chi)(\lambda_k), \\ e_{k+1}(\Psi) &= \sigma_{k+1}(\Psi)(\lambda_{k+1}).\end{aligned}$$

We consider $\bar{\gamma}_2$, and thus $\bar{\tau}_{i,j}$ for $0 \leq i < j \leq n$, to be fixed; so that increasing or decreasing angles φ , χ , or Ψ will be by moving the σ_i . Let $(-\sigma_0(\varphi_1), \bar{\tau}_{0,k}, \sigma_k(\chi_1))$ be the triangle in \bar{M} which corresponds to $(-\sigma_0, \tau_{0,k}, \sigma_k)$. Observe that $\varphi_1 < \varphi_0$ and $\chi_1 > \chi_0$ by the induction hypothesis. It may help to refer to Fig. 8. Then again, it may not.

Note that in Theorem 4.2.2 $(-\sigma_0(\varphi_1), \tau_{0,k}, \sigma_k(\chi_1))$ is contained in \bar{M}_+ , while in Theorem 4.2.1 we have no such assurance. This is one of the reasons why in Theorem 4.2.1 it is convenient to work in \bar{M}^* .

Since Theorem 4.2.1 will be used as a lemma for Theorem 4.2.2, it is necessary to first finish Theorem 4.2.1, and then come back to Theorem 4.2.2. Steps (7) and (8) refer only to Theorem 4.2.1.

(7) We will first show that as φ is decreased from φ_0 to φ_1 , the distance from $\sigma_0(\varphi)$ to \bar{x}_{k+1}^* is decreased. Specifically, since $\sigma_0(\varphi_0) \subset \bar{M}_+^* \cup \bar{p}^*$, Lemma 3.3.7 implies that $(d/d\varphi)d(e_0(\varphi), \bar{x}_{k+1}^*)|_{\varphi=\varphi_0} \geq 0$. This remains true, as φ is decreased, so long as

$$\sigma_0(\varphi) \subset (\bar{M}_+^* \cup T_{d(\bar{x}_{k+1}^*, \bar{p}^*) + \delta_1}(\bar{x}_{k+1}^*)),$$

and $\sigma_0(\varphi)$ is either in \bar{M}_+^* or below \bar{x}_{k+1}^* . The latter condition clearly remains in force as φ is decreased. To see that the former does also, suppose that

$$e_0(\bar{\varphi}) \in \text{Cl}(T_{d(\bar{x}_{k+1}^*, \bar{p}^*)}(\bar{x}_{k+1}^*))$$

for some $\bar{\varphi}$. But since the hypotheses of Lemma 3.3.7 are satisfied, we see that $d(e_0(\varphi), \bar{x}_{k+1}^*)$ continues to decrease, which forces $e_0(\varphi)$ to remain in $T_{d(\bar{x}_{k+1}^*, \bar{p}^*) + \delta_1}(\bar{x}_{k+1}^*)$. This of course applies as well to each point on $\sigma_0(\varphi)$, so that

$$d(e_0(\varphi_1), \bar{x}_{k+1}^*) < d(e_0(\varphi_0), \bar{x}_{k+1}^*) = d(\bar{e}_1^*, \bar{x}_{k+1}^*).$$

This same argument applied to $\sigma_k(\chi)$ implies that

$$d(e_0(\varphi_1), \bar{x}_{k+1}^*) = d(e_k(\chi_1), \bar{x}_{k+1}^*) \geq d(e_k(\chi_0), \bar{x}_{k+1}^*).$$

Now just as in step (4) we conclude that

$$d(e_k(\chi_0), \bar{x}_{k+1}^*) > L[\sigma_{k+1}] = d(e_1, x_{k+1}).$$

Combining all inequalities yields

$$d(\bar{e}_1^*, \bar{x}_{k+1}^*) \geq d(e_1, x_{k+1}),$$

which is (A).

(8) We will now show that (B) holds for $(\sigma_0, \tau_{0,k+1}, \sigma_{k+1})$.

The techniques of step (5) show that (B) is true for the triangle $(-\sigma_k, \tau_{k,k+1}, \sigma_{k+1})$. Let $(\sigma_k(\chi_3), \tau_{k,k+1}^*, \sigma_{k+1}(\Psi_3))$ denote the corresponding triangle in \bar{M}^* . Note that $\Psi_3 > \Psi_0$, and that increasing Ψ_3 will bring $e_{k+1}(\Psi)$ still closer to \bar{x}_k^* .

Decrease φ_1 to φ_2 , so that $d(e_0(\varphi_2), \bar{x}_{k+1}^*) = L[\sigma_{k+1}]$. Clearly Lemma 3.3.7 continues to apply during this movement. Let Ψ_2 be the angle determined by the (unique) geodesic $\sigma_{k+1}(\Psi_2) \in \text{Geo}(\bar{x}_{k+1}^*, e_0(\varphi_2))$.

As noted above, $\varphi_2 < \varphi_1 < \varphi_0$. To see that $\Psi_2 > \Psi_0$, observe that $d(e_0(\varphi_2), \bar{x}_k^*) < L[\sigma_k]$, and thus (working in \bar{M} if needed) Ψ_3 must be *increased* to bring $e_{k+1}(\Psi)$ nearer to \bar{x}_k^* ; i.e., so that $d(e_{k+1}(\Psi_2), \bar{x}_k^*) < L[\sigma_k]$.

Hence B is true for $(\sigma_0, \tau_{0,k+1}, \sigma_{k+1})$, and the proof of Theorem 4.2.1 is complete.

We now repeat steps (7) and (8) for Theorem 4.2.2, using Theorem 4.2.1 as needed.

(9) Since $(-\sigma_0(\varphi_1), \bar{\tau}_{0,k}, \sigma_k(\chi_1)) \subset \bar{M}_+$ by the induction hypothesis, it is clear that $d(\bar{e}_1, \bar{x}_{k+1}) \geq d(e_k(\chi_1), x_{k+1})$. Since $\sigma_k(\chi_0) \subset \bar{M}_+$, it is clear that $d(e_k(\chi_1), \bar{x}_{k+1}) \geq d(e_k(\chi_0), \bar{x}_{k+1})$. Theorem 4.2.1, applied to the hinge $(\tau_{0,k}, \sigma_k, \beta_k)$, implies that $d(\sigma_k(t), p) \leq d(\bar{\sigma}_k(t), \bar{p})$. Thus Lemma 4.1.5 implies that $d(e_k(\chi_0), \bar{x}_{k+1}) \geq d(e_1, x_{k+1})$. Hence $d(\bar{e}_1, \bar{x}_{k+1}) \geq d(e_1, x_{k+1})$.

(10) We again apply Theorem 4.2.2, this time to the hinge $(\tau_{0,k+1}, \sigma_{k+1}, \beta_{k+1})$, and conclude that $d(\sigma_{k+1}(t), p) \leq d(\bar{\sigma}_{k+1}(t), \bar{p})$. Thus Lemma 4.1.5 implies that $d(e_{k+1}(\Psi_0), \bar{x}_k) \geq L[\sigma_k]$. (B) now follows for the triangle $(-\sigma_0, \tau_{0,k+1}, \sigma_{k+1})$ in a manner analogous to that in step (8). Uniqueness is discussed in step (11).

(11) Loose ends: First, the uniqueness of the triangle in (B) of Theorem 4.2.2 is clear since the triangle is contained in \bar{M}_+ and has one vertex at p .

For the assumption in step (3), note that any flattening surface of revolution is a deformation of \mathbf{R}^2 through flattening surfaces. Suppose that $\{S^i\}$ is such a deformation with $S^0 = \mathbf{R}^2$ and $S^1 = \bar{M}$, and let γ_i' and σ_i' denote the geodesics on S^i resulting from the construction in step (2). Let B^i denote the

region of S_+^t which is bounded by $\mu_{e_1}^t$ and $\mu_{e_2}^t$. Steps (4)–(10) show that if σ_i^t is in S_+^t , it is actually restricted to B^t . Since the σ_i^t are continuous with respect to t , since no σ_i^t passes through p , and since the σ_i^0 are clearly in S_+^0 , it follows that the σ_i^1 are in $S_+^1 = \overline{M}_+$. Here we are also using the fact that $\alpha_0 < \pi$.

For the assumption in step (1) concerning δ -correspondence, construct \overline{M} corresponding to M . Since the theorem holds if the metric of \overline{M} is multiplied by any constant greater than 1, it must, by continuity, hold on \overline{M} itself.

If for any x , $\langle \gamma_x', \gamma_2' \rangle(x) = -1$ (see step (1)), then $\alpha_0 = \pi$, and the theorem holds by continuity.

Again referring to the assumptions in step (1), suppose that $\bar{t} \in [0, b_2]$ is the smallest number such that $\langle \gamma_y', \gamma_2' \rangle(y) = +1$, where $y = \gamma_2(\bar{t})$. Then, by continuity, the theorem holds for the hinge $(\gamma_1, \gamma_2][0, \bar{t}]$, α_0 . since $\gamma_2][\bar{t}, b_2]$ is contained in the minimal geodesic from y to e_1 , the theorem clearly holds for $(\gamma_1, \gamma_2, \alpha)$.

5. MANIFOLDS DIFFEOMORPHIC TO EUCLIDEAN SPACE

Conjecture (Cheeger and Gromoll). *If M^n is a complete open nonnegatively curved manifold, and all sectional curvatures at some point $x \in M$ are positive, then M is diffeomorphic to \mathbf{R}^n .*

This conjecture is known to be true if $n = 2, 3$ or 4 ; see [5], [3] and [4] respectively. If the words "some point" are replaced by "every point", the conjecture follows from [9] or [3].

In this section we apply Theorem 4.2 to the soul construction of Cheeger and Gromoll [3] to prove a weak version of this conjecture. In dimensions greater than 4, this is the first such result which does not require M to have positive curvature everywhere off a compact set.

5.1. The soul of a manifold

We begin by recalling several facts concerning the above mentioned soul construction. M^n will continue to denote a complete open nonnegatively curved manifold.

5.1.1. Definition. A set $C \subset M$ is said to be *totally convex* if for any $x, y \in C$, $\text{Geo}(x, y) \subset \text{Geo}(C)$.

5.1.2. Theorem (Cheeger and Gromoll). *M contains a compact totally geodesic submanifold S without boundary which is totally convex, $0 \leq \dim S < \dim M$.*

Proof. We merely outline part of the construction of the set S . For further details see Cheeger and Gromoll [3]; or Cheeger and Ebin [2].

Let $x \in M$ and let $\gamma : [0, \infty) \rightarrow M$ be any geodesic ray beginning at x ; i.e., γ is a globally minimal geodesic. Let $B_x(\gamma) = \{y \in M \mid d(\gamma(t), y) < t, t \in [0, \infty)\}$, and set $H_x(\gamma) = M \setminus B_x(\gamma)$. We call $H_x(\gamma)$ a *complementary half space*.

Define $C_x = \bigcap (H_x(\gamma))$ as γ ranges over all rays beginning at x . It is not difficult to show that C_x is compact. S is now constructed as a certain subset of C_x .

5.1.3. Definition. The set $S \subset M$ constructed in Theorem 5.1.2 is called a *soul of M* .

5.1.4. Theorem (Cheeger and Gromoll). *Let S be a soul of M . Then M is diffeomorphic to the normal bundle $\nu(S)$ of S in M . Furthermore, if $K(x) > 0$ for each $x \in S$, then S is a point, and hence M is diffeomorphic to \mathbf{R}^n .*

Proof. See Cheeger and Gromoll [3] and either Poor [16] or Šarafutdinov [17].

5.2. Compact half spaces and shriveled souls

Theorem. *Let M^n be a complete open nonnegatively curved manifold. Suppose $p \in M$ and γ is a ray starting at p .*

5.2.1. *If there is an $r_1 \in \mathbf{R}$ such that for each $x \in T_{r_1}(p)$, $K(x) > (\pi/3r_1)^2$, then $H_x(\gamma)$ is compact.*

5.2.2. *If there is an $r_2 \in \mathbf{R}$ such that for each $x \in T_{r_2}(p)$, $K(x) > (\pi/\lambda r_2)^2$, where $\lambda \cong 2.46057$, then M is diffeomorphic to \mathbf{R}^n .*

Proof. This is an easy application of Theorem 4.2.2 to the construction outlined in Theorem 5.1.2. We first need to construct appropriate comparison surfaces.

Let C_r^H denote a cone with spherical cap of curvature H and radius r . That is, C_r^H is the surface of revolution generated by the curve $f : \mathbf{R}_+ \rightarrow \mathbf{R}^2$, f defined by

$$f(t) = \begin{cases} \left(\frac{\sin(t)}{\sqrt{H}}, \frac{1 - \cos(t)}{\sqrt{H}} \right), & 0 \leq t \leq \theta, \\ \left(\frac{\sin(\theta)}{\sqrt{H}} + (t - r) \sin(\theta), \frac{1 - \cos(\theta)}{\sqrt{H}} + (t - r) \cos(\theta) \right), & t > \theta, \end{cases}$$

where $\theta = r\sqrt{H}$. See Fig. 9 (a).

If $x \in C_r^H$ and $y \in \bar{\mu}_x$, let g denote a minimal geodesic from x to y . Let P denote the parallel through the point $(H^{-1/2} \cdot \sin(\theta), H^{-1/2} \cdot (1 - \cos(\theta))) \in f \subset S$, and let $p = d(x, P)$, $q = d(y, P)$. Let s and φ be as indicated in Fig. 9 (b).

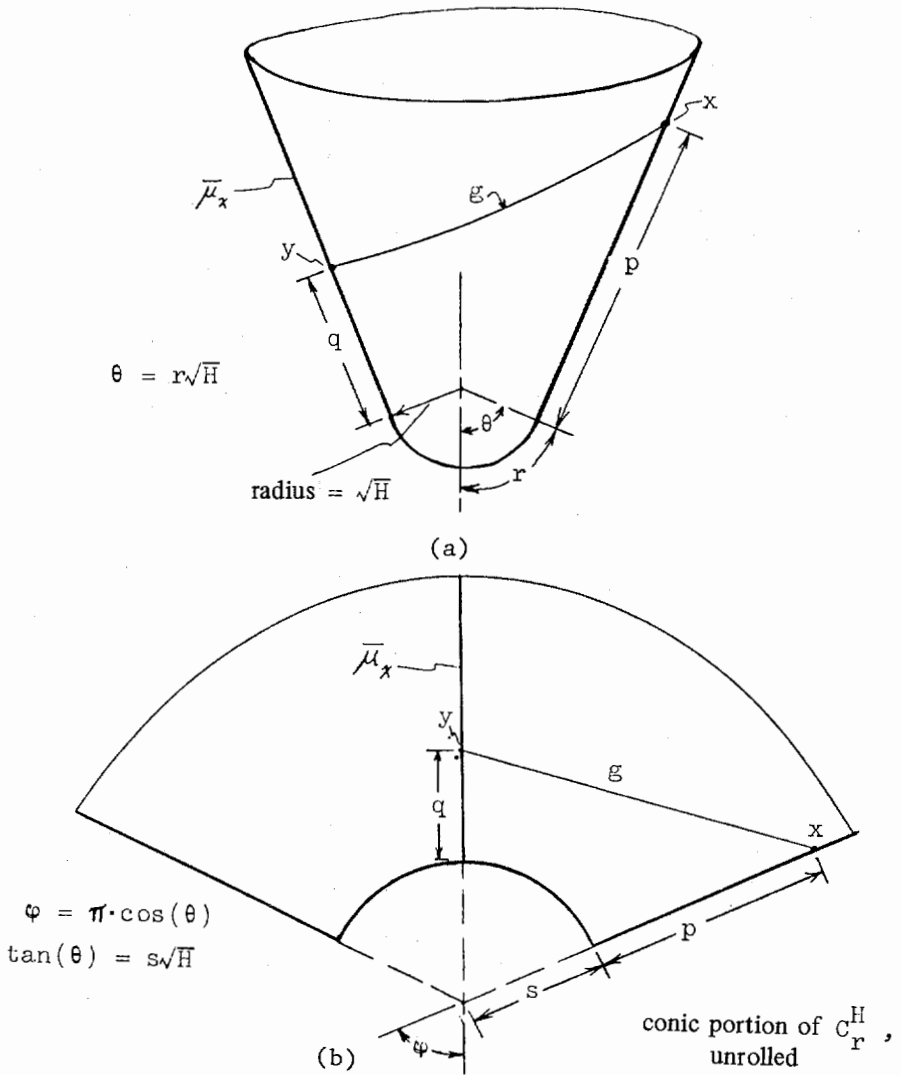


FIG. 9

We will now find conditions on θ which allow us to choose $x \in C_r^H$ so that $L[g] \leq p + r$ in the cases

(1) g minimal between x and $\bar{\mu}_x$, and

(2) $q = 0$.

Case (1). $L[g] = (p + s) \sin(\varphi)$.

We need conditions so that

$$(p + s)\sin(\varphi) < (p + r),$$

or equivalently,

$$p(1 - \sin(\varphi)) + (r - s \cdot \sin(\varphi)) > 0.$$

Hence if $\sin(\varphi) \neq 1$, we may choose p large enough to insure that the inequality holds. Thus we need $\varphi < \pi/2$, or $\pi/3 < \theta \leq \pi/2$; which is to say

$$\frac{\pi}{3\sqrt{H}} < r \leq \frac{\pi}{2\sqrt{H}}.$$

Case (2). $L[g]^2 = s^2 = (p + s)^2 - 2s(p + s) \cdot \cos(\varphi)$.

We need conditions so that

$$s^2 + (p + s)^2 - 2s(p + s) \cdot \cos(\varphi) < (p + r)^2,$$

i.e.,

$$2p[s \cdot \cos(\varphi) + r - s] + [r^2 + 2s^2 \cdot \cos(\varphi) - 2s^2] > 0.$$

Hence if $s \cdot \cos(\varphi) + r - s > 0$, we may choose p large enough so that the inequality holds.

Thus the condition is

$$\frac{\tan(\theta)}{\sqrt{H}} \cdot \cos(\varphi) + \frac{\theta}{\sqrt{H}} - \frac{\tan(\theta)}{\sqrt{H}} > 0,$$

i.e.,

$$[\tan(\theta)] \cdot [\cos(\pi \cos(\theta))] + \theta - \tan(\theta) > 0.$$

This is true for $\pi/\lambda < \theta \leq \pi/2$, where $\lambda \cong 2.46057$. Thus we need

$$\frac{\pi}{\lambda\sqrt{H}} < r \leq \frac{\pi}{2\sqrt{H}}.$$

Despite the fact that C_r^H is not smooth, we may use it as a comparison surface by virtue of the previously mentioned approximation theorems in [1].

For Theorem 5.2.1, choose $H_1 \in \mathbf{R}$ such that

$$(\pi/3r_1)^2 \leq H_1 \leq K(x) \quad \forall x \in T_{r_1}(p),$$

and for Theorem 5.2.2 choose $H_2 \in \mathbf{R}$ such that

$$(\pi/\lambda r_2)^2 \leq H_2 \leq K(x) \quad \forall x \in T_{r_2}(p).$$

Let η be any minimal geodesic in M which starts at p .

In Theorem 5.2.1, by comparison with $C_{r_1}^{H_1}$, there is some $t_0 > 0$ such that $\eta(t) \in B_p(\gamma)$ for $t > t_0$; and thus $H_p(\gamma) \subset T_{t_0}(p)$.

In Theorem 5.2.2, by comparison with $C_{r_2}^{H_2}$, $\eta(t) \in B_p(\gamma)$ for $t > r_2$. Thus $S \subset H_p(\gamma) \subset T_{r_2}(p)$. But $K(x) > 0$ for all x in $T_{r_2}(p)$. Hence by Theorem 5.1.4, M is diffeomorphic to \mathbf{R}^n .

References

- [1] A. D. Alexandrov & V. A. Zalgaller, *Intrinsic geometry of surfaces*, Transl. Math. Monographs, Vol. 15, Amer. Math. Soc., 1967.
- [2] J. Cheeger & D. Ebin, *Comparison theorems in Riemannian geometry*, North Holland/American Elsevier, Amsterdam-Oxford-New York, 1975.
- [3] J. Cheeger & D. Gromoll, *On the structure of complete manifolds of nonnegative curvature*, Ann. of Math. **96** (1972) 413–443.
- [4] ———, private communication.
- [5] S. Cohn-Vossen, *Kürzeste Wege und Totalkrümmung auf Flächen*, Compositio Math. **2** (1935) 69–133.
- [6] ———, *Totalkrümmung und geodätische Linien auf einfachzusammenhängenden offen vollständigen Flächenstücken*, Rec. Math. Moscow **43** (1936) 139–163.
- [7] G. Darboux, *Leçons sur la théorie générale des surfaces*, Livre VI, Gauthier-Villars et Fils, Paris, 1894.
- [8] H. Gluck & D. Singer, *The existence of nontriangulable cut loci*, Bull. Amer. Math. Soc. **82** (1976) 599–602.
- [9] D. Gromoll & W. Meyer, *On complete open manifolds of positive curvature*, Ann. of Math. **90** (1969) 75–90.
- [10] W. Klingenberg, *Eine Vorlesung über Differentialgeometrie*, Springer, Berlin, 1973.
- [11] D. Ling, *Geodesics on surfaces of revolution*, Trans. Amer. Math. Soc. **59** (1946) 415–429.
- [12] H. von Mangoldt, *Über diejenigen Punkte auf positive gekrümmten Flächen, welche die Eigenschaft haben, dass die von Ihnen ausgehenden geodätischen Linien nie aufhören, kürzeste Linien zu sein*, J. Reine Angew. Math. **91** (1881) 23–53.
- [13] S. B. Meyers, *Connection between differential geometry and topology*. I, Duke Math. J. **1** (1935) 376–391.
- [14] ———, *Connections between differential geometry and topology*. II, Duke Math. J. **2** (1936) 95–102.
- [15] H. Poincaré, *Sur les lignes géodesiques des surfaces convexes*, Trans. Amer. Math. Soc. **6** (1905) 237–274.
- [16] W. A. Poor, *Some results on nonnegatively curved manifolds*, J. Differential Geometry **9** (1974) 583–600.
- [17] V. A. Sarafutdinov, *Complete open manifolds of nonnegative curvature*, Siberian Math. J. **15** (1974) 126–136.
- [18] V. A. Toponogov, *Riemann spaces with the curvature bounded below*, Uspehi Mat. Nauk **14** (1959) 87–130, in Russian.

HUGHES AIRCRAFT COMPANY, CULVER CITY, CA