

Thin Riemannian manifolds with boundary

Stephanie B. Alexander, Richard L. Bishop

University of Illinois at Urbana-Champaign, 1409 W. Green, Urbana, IL 61801, USA
 (e-mails: sba@math.uiuc.edu; bishop@math.uiuc.edu)

Received: 27 December 1995 / Accepted: 23 August 1996

Mathematics Subject Classification (1991): 53C20, 53C29

1. Introduction

Let M be a complete connected Riemannian manifold with boundary B , K be the sectional curvature of the interior, and κ be the normal curvature of B . The *inradius* of M is

$$\text{inr} = \sup\{d(p, B) : p \in M\}.$$

Thus for every $r < \text{inr}$ there is some metric ball of radius r in M that does not intersect B . By a *thin* manifold with boundary M we mean one for which the scale-free invariant

$$\text{inr}^2 \max\{\sup |K|, \sup \kappa^2\}$$

is small. The main purpose of this paper is to prove that a sufficiently thin manifold with boundary is fibered over a manifold without boundary:

Theorem 1.1. *There is a dimension-independent constant c ($\geq .075$) such that if M satisfies $\text{inr}^2 \max\{\sup |K|, \sup \kappa^2\} < c^2$, then either M is diffeomorphic to the product of a manifold without boundary and an interval or M can be doubly covered by such a product.*

This theorem, but with a dimension-dependent constant $c_n \leq n^{-n}$, was stated without proof by Gromov in his 1978 address to the International Congress [G, Theorem 5].

We also show that a complete, connected, simply connected Riemannian manifold X without boundary is separated by any proper connected hypersurface H into components that cannot be small relative to the curvature of X and the normal curvature of H . More generally:

Theorem 1.2. *There is a universal constant c ($\geq .075$) such that if X is a complete connected Riemannian manifold without boundary with $|K| \leq 1$ and $H_1(X; \mathbb{Z}/2\mathbb{Z}) = 0$, and H is a smooth connected hypersurface properly imbedded in X with normal curvatures at most 1, then the closure of each component of $X - H$ has inradius at least c .*

By a theorem of Lagunov, if H is a compact connected hypersurface in E^n , $n \geq 3$, with normal curvatures at most 1, then the bounded component of the complement of H contains a ball of radius at least $((2/\sqrt{3}) - 1)$ and this estimate is best possible [L1], [L2]. Lagunov and Fet showed that the sharp bound increases to $(\sqrt{3}/2 - 1)$ under assumptions of simple topology [LF1], [LF2]. (The situation is radically different for $n = 2$, where the bound is 1 by a theorem of Pestov and Ionin [PI]; also see [HT].) It follows from Lagunov's example that the optimal constant c in Theorems 1.1 and 1.2 is at most $((2/\sqrt{3}) - 1) \sim .15$. A similar example based on hyperbolic geometry shows that c can be at most $\log(2/\sqrt{3}) \sim .144$.

As for metric structure in Theorem 1.1, we remark below that M or its double cover need not be close to a metric product for c (or c_n) sufficiently small. However, this holds if there is a uniform bound on diam/inr , and in general M is Lipschitz close to a manifold with boundary imbedded in a metric product of the same dimension (see Sect. 5).

Recall that p is a *cut point* of the boundary B if there is a geodesic segment σ realizing $d(p, B)$, but no extension of σ realizes distance to B . The *degree* of p is the number of such minimizing segments; a *footpoint* of p is the endpoint on B of such a segment. Theorem 1.1 will be proved by showing that if c is sufficiently small, then every cut point of B has degree 2. More generally, it follows from our argument that there is an increasing sequence of positive constants $c(k)$, $k = 2, 3, \dots$, such that if the curvature-normalized inradius of M is less than $c(k)$, then the cut points of B have degree at most k .

Section 2 establishes a critical radius for a cut point p : out to this radius, a boundary geodesic starting from a footpoint of p will remain visible from p through the interior of M , without boundary interference, and will recede from p . Thus it becomes profitable to consider the manifold with boundary obtained by lifting to $T_p M$ the geodesics from p whose relative interiors lie in the interior of M and whose lengths are less than the critical radius. Our program for bounding the degree of p is to show that *in the boundary of the lift*, the footpoints of p have injectivity radius uniformly bounded below; and for c sufficiently small, the corresponding normal coordinate neighborhoods subtend large disjoint cones with vertex p . Our methods are those of comparison geometry; for the injectivity radius bound, they build on previous work of the authors [ABB], [AB]. Section 3 gives a construction that reduces Theorem 1.1 by comparison arguments to a model space problem. Section 4 carries out estimates solving this problem.

Below we shall rescale so that $|K| \leq 1$ and $|\kappa| \leq 1$. We denote by M_δ ($\delta = \pm$) a model manifold with boundary, either a closed disk of diameter $\pi/2$ in the 2-sphere of curvature 1 ($\delta = +$) or the closure of the outside of a horocycle in

the hyperbolic plane of curvature -1 ($\delta = -$). Note that the boundary B_δ of M_δ has signed curvature 1 ($\delta = +$) or -1 ($\delta = -$). The relative comparison lemmas in the following section have independent interest for the study of manifolds with boundary. Although they are simplified by our curvature normalizations, it will be clear how to generalize them to arbitrary curvature bounds.

The authors thank Albrecht Dold and Anton Petrunin for clarifying conversations. They are also grateful to the referee for a careful reading and helpful suggestions.

2. Comparisons of visual distance and angle

Throughout this section, M will denote a complete manifold with boundary B , having $|K| \leq 1$ and $|\kappa| \leq 1$. Under these curvature assumptions, the conjugate distance of the interior of M is at least $\pi/2$ and the focal distance of B in M is at least $\pi/4$. Let p be a point of M whose distance from B is r_0 , where r_0 is less than the focal distance of B in M . We denote by $t(r_0)$ the length of a geodesic in the model space M_- that starts at a distance r_0 from the boundary and ends tangent to it, touching the boundary only at its endpoint. Then $r_0 = \log \cosh t(r_0)$.

Lemma 2.1. *Let p be a point of M whose distance from B is r_0 , where r_0 is less than the focal distance of B in M . Then a geodesic from p of length less than $t(r_0)$ cannot be tangent to B .*

The idea here is that, given a boundary tangency closer to p than $t(r_0)$, we could force a boundary geodesic closer to p than the allowable distance r_0 ; see the proof below.

Henceforth we assume $r_0 < \pi/4$.

We say that a point of M is *visible from p* if it is the endpoint of a geodesic from p with relative interior in the interior of M (a *line of visibility*). These lines of visibility are not required to be minimizers. In this section we study the part of M visible from p along geodesics of length less than $d(r_0)$, where

$$d(r_0) = \min\left\{t(r_0), \frac{\pi}{2} - r_0\right\}. \quad (1)$$

The lines of visibility from p of length less than $d(r_0)$ lift via the exponential map at p to a smooth manifold with boundary $M^* = M^*(p, r_0)$ in $T_p M$ carrying the pullback Riemannian metric, as follows from Lemma 2.1 and the fact that $d(r_0)$ is less than the conjugate distance of M . Thus \exp_p acts on the boundary $B^* = B^*(p, r_0)$ of M^* as a local isometry, not necessarily injective, into B .

Now consider a geodesic γ of B^* . Let s denote the arclength parameter of γ ; r , its radial distance from the origin, which we again denote by p ; and φ , the angle γ makes with the outward radial direction from p . For comparison purposes, consider a point p_δ in M_δ and a model curve γ_δ in the boundary of M_δ having arclength parameter s_δ and the same initial conditions as γ . Namely, at $s_\delta = 0$, its distance r_δ from p_δ agrees with $r(0)$, and the angle φ_δ it makes with the outward radial direction from p_δ agrees with $\varphi(0)$.

Lemma 2.2. *A geodesic of $B^* = B^*(p, r_0)$ which starts at a footpoint of p on B^* always recedes from p and may be reparametrized by its radial distance r from p . Furthermore,*

- (i) $\varphi_-(r) \leq \varphi(r) \leq \varphi_+(r)$,
- (ii) $s_-(r) \leq s(r) \leq s_+(r)$.

Proof of Lemmas 2.1 and 2.2. First let γ be any geodesic of B that is visible from p and on whose relative interior the angle φ with the outward radial direction lies in $(0, \pi/2)$. By first variation of arclength, $dr/ds = \cos \varphi > 0$, so γ may be reparametrized by the radial distance r from p . We begin by proving that γ satisfies the lefthand inequalities of (i) and (ii) if $r \leq t(r_0)$, and the righthand inequalities if $r \leq \frac{\pi}{2} - r_0$. (Below, when we say that (i) and (ii) are satisfied, these restrictions will be understood. This is the motivation for the definition of $d(r_0)$ in (1).)

We work in the ruled surface S whose generators are the lines of visibility from p to γ . The geodesic curvature κ_S of γ in S is related to its curvature κ in M by $|\kappa_S| \leq |\kappa| \leq 1$. Furthermore, while S need not inherit the lower sectional curvature bound of M , the orthogonal S -Jacobi field along any generator is also an M -Jacobi field. Therefore by the Rauch comparison theorem, the logarithmic derivative $f = g_r/g$ of its length g satisfies

$$\cot r \leq f(r) \leq \coth r. \quad (2)$$

Now observe that by expressing κ_S in terms of the parameter r one obtains a linear equation in $\sin \varphi$. Specifically, write

$$\gamma' = w(\cos \varphi E + \sin \varphi F),$$

where E, F are orthonormal, E points in the outward radial direction, and $w = \|\gamma'\| = 1/\cos \varphi$. Let Y be a parallel vector field in S along γ , ψ be the angle from Y to γ , and θ be the angle from Y to F . Hence, $\psi = \theta - \frac{\pi}{2} + \varphi$ and $\kappa_S = \psi'/w = (\varphi' + \theta')/w$. But θ'/w is the rate at which F rotates when moving along any curve with speed w ; here it can be rewritten as $\theta'/w = g'/w = f \sin \varphi$, the rate at which E rotates when moving in the direction of γ' . Therefore

$$\kappa_S = (\sin \varphi)' + f \sin \varphi. \quad (3)$$

To estimate $\sin \varphi$, multiply (3) by $\exp(\int_{r(0)}^r f(u) du)$ and integrate:

$$\begin{aligned} \sin \varphi \cdot \exp\left(\int_{r(0)}^r f(u) du\right) &= \sin \varphi(0) + \int_{r(0)}^r \kappa_S(t) \exp\left(\int_{r(0)}^t f(u) du\right) dt \quad (4) \\ &\geq \sin \varphi(0) - \int_{r(0)}^r \exp\left(\int_{r(0)}^t f(u) du\right) dt \\ &\geq \sin \varphi(0) - \frac{\cosh r - \cosh r(0)}{\sinh r(0)}, \end{aligned}$$

where the last step is directly from the bound $f(u) \leq \coth u$ in (2). Multiplying by $\exp(-\int_{r(0)}^r f(u) du) \geq (\sinh r(0))/\sinh r$ gives

$$\sin \varphi \geq \frac{\sin \varphi(0) \cdot \sinh r(0) - (\cosh r - \cosh r(0))}{\sinh r}. \quad (5)$$

Since the curve γ_- in M_- satisfies equality at each step, the right side of (5) is $\sin \varphi_-(r)$ for $r \leq t(r_0)$ and hence γ satisfies the left side of (i).

In a similar fashion we obtain an upper bound on $\sin \varphi$. Starting with (4), we have

$$\sin \varphi \cdot \exp\left(\int_{r(0)}^r f(u) du\right) \leq \sin \varphi(0) + \int_{r(0)}^r \exp\left(\int_{r(0)}^t f(u) du\right) dt.$$

Multiplying both sides by $\exp(-\int_{r(0)}^r f(u) du)$ makes the second term on the right $\int_{r(0)}^r \exp(-\int_t^r f(u) du) dt$. Replacing f by its lower bound in (2), we conclude that γ satisfies the right side of (i). Here the restriction $r \leq \frac{\pi}{2} - r_0$ is imposed by the range of r on the model curve, which lies in the boundary circle of diameter $\pi/2$ in M_+ . Note that (i) implies $ds_-/dr \leq ds/dr = \sec \varphi \leq ds_+/dr$, and hence (ii) follows immediately.

Now suppose a visual tangency to B occurred closer to p than $t(r_0)$. Let $q \in B$ be a point of visual tangency closest to p , and consider the B -geodesic having the outward radial direction at q as velocity. Running backward from q along this geodesic, the choice of q guarantees that we remain visible from p for as long as r decreases. Thus there is an open subarc γ for which $0 < dr/ds = \cos \varphi < 1$ and $\varphi = 0$ at the right endpoint. Moreover, γ may be chosen so that φ is arbitrarily close to $\pi/2$ at the left endpoint, since $\varphi = 0$ cannot occur to the left of q and dr/ds cannot be bounded away from 0 indefinitely. As we have just seen, γ satisfies the lefthand inequality in (i). Therefore the value of r at the right endpoint of γ is at least $t(\tilde{r}_0)$, where \tilde{r}_0 denotes the limit of r as the left endpoint moves so that $\varphi \rightarrow \pi/2$. By hypothesis, $t(\tilde{r}_0) < t(r_0)$. But then $\tilde{r}_0 < r_0$, contradicting the definition of r_0 and proving Lemma 2.1.

Finally consider a B -geodesic γ^* that is visible from p along geodesics of length less than $d(r_0)$ and whose starting point q realizes the distance r_0 from p to B . By Lemma 2.1, since $d(r_0) \leq t(r_0)$ then $\varphi \neq 0$ on γ^* . Since r_0 is less than the focal distance of B , we have $(d^2r/ds^2)(0) > 0$. Therefore $0 < dr/ds = \cos \varphi < 1$ on an open subarc γ of γ^* with q as left endpoint. We have shown above that (i) and (ii) hold for such subarcs γ . Applying the righthand side of (i) to these subarcs, we see that since $d(r_0) \leq \frac{\pi}{2} - r_0$ then $\varphi \neq \pi/2$ on γ^* . This is because the boundary circle B_+ of M_+ has diameter $\pi/2$; therefore φ_+ , the angle between B_+ and the radial direction from a point at distance r_0 from B_+ , is less than $\pi/2$ for $r_0 < r < \frac{\pi}{2} - r_0$. Hence (i) and (ii) hold for γ^* . This completes the proof of Lemma 2.2. \square

Next consider the metric completion \bar{M}^* of $M^* = M^*(p, r_0)$ under the metric of infimum of pathlengths. This *visual lift* \bar{M}^* is homeomorphic to the closure

of M^* in $T_p M$, and is the lift of the lines of visibility from p in M of length less than or equal to $d(r_0)$. The boundary points of \bar{M}^* are either in B^* or at distance exactly $d(r_0)$ from p . Since \bar{M}^* is compact, it is a *geodesic* metric space in the sense that any two points are joined by a minimizer. Note that the geodesics of \bar{M}^* , where they run along B^* , can bend relative to the geometry of the interior and bifurcate.

Recall that a geodesic metric space has *curvature* $\leq K$ in the sense of Alexandrov if every point has a neighborhood in which for any minimizing geodesic triangle with vertices in the neighborhood and perimeter less than $2\pi/\sqrt{K}$ ($= \infty$ if $K \leq 0$), the distance between points on the triangle is at most equal to the distance between corresponding points on the comparison triangle with the same sidelengths in the model space S_K . Here, S_K denotes the 2-dimensional sphere, Euclidean or hyperbolic space of constant curvature K . A *CAT(K)-space* satisfies in addition the condition that minimizers of length $< \pi/\sqrt{K}$ are unique and vary continuously with their endpoints. It is equivalent to say that triangle comparisons hold globally as well as locally, that is, they hold for all minimizing geodesic triangles of perimeter less than $2\pi/\sqrt{K}$ [Av1], [Av2]. The next theorem summarizes several results about geodesic metric spaces that will be used in this paper; part (ii) was proved for Riemannian spaces in [BK, p. 100], [GLP, p. 188], [J].

Theorem 2.3.

- (i) [ABB] *A complete Riemannian manifold with boundary, regarded as a geodesic metric space, has curvature $\leq K$ in the sense of Alexandrov if and only if its interior sectional curvatures and its boundary sectional curvatures on outward-pointing boundary sections are at most K .*
- (ii) [AB] *In a complete geodesic metric space with curvature $\leq K$, suppose D is an open metric ball of radius at most $\pi/2\sqrt{K}$, in which minimizers from the centerpoint are unique and vary continuously with their right endpoints. Then D is a CAT(K)-space. Also, the radial distance from the centerpoint is convex on D and strictly convex on nonradial geodesics.*
- (iii) [AB] *In a CAT(K)-space, a closed, nonconstant curve of geodesic curvature at most k is no shorter than a complete circle of curvature k in S_K . (The curve is allowed to join nonsmoothly at the terminus.)*

Remark. The geodesic curvature of a curve in a metric space is defined as in [AB], using arc/chord comparisons to curves of constant curvature in S_K . It agrees with the usual notion of geodesic curvature when applied to a smooth curve in a Riemannian manifold.

Lemma 2.4. *The visual lift $\bar{M}^* = \bar{M}^*(p, r_0)$, regarded as a geodesic metric space, has the following properties:*

- (i) *Except possibly at its endpoints, a minimizer in \bar{M}^* lies entirely in the manifold with boundary M^* .*
- (ii) *\bar{M}^* is a CAT(2)-space.*

Proof. (i). First note that the uniqueness of minimizers from p in \bar{M}^* and their continuous dependence on their righthand endpoints is an automatic consequence of the exponential lifting that defines \bar{M}^* . Let β be a minimizer in \bar{M}^* . We must show that the distance from p to a relative interior point of β is strictly less than d_0 . Without loss of generality, it may be assumed that β contains no nontrivial segment of a minimizer from p . Consider the ruled surface S in \bar{M}^* whose generators are the minimizers from p to β . By the construction of \bar{M}^* , it is easy to see that the Riemannian metric of its interior may be extended across its boundary, and hence S may be realized as a ruled surface in a Riemannian manifold without boundary of curvature $\leq (1 + \epsilon)$. Since β is continuous and the minimizers from p vary continuously with their righthand endpoints, we may invoke the theorem of Alexandrov stating that ruled surfaces in metric spaces of curvature $\leq K$ inherit the same curvature bound [Av3]. Therefore S has curvature $\leq (1 + \epsilon)$. Regarding S as an open metric ball in itself of radius $\pi/2(1+\epsilon)^{1/2} > d(r_0)$ and applying Theorem 2.3(ii), we conclude that the distance from p to β is strictly convex.

(ii). Note that under our curvature bounds, the sectional curvatures of B are at most 2 by the Gauss equation. Given part (i), the proof that \bar{M}^* has curvature ≤ 2 in the sense of Alexandrov follows without essential change the corresponding proof in [ABB, Sect. 2.3] (see Theorem 2.3 (i) above). To show that \bar{M}^* is a CAT(2)-space, by Theorem 2.3 (ii) it is only necessary to check that $d(r_0) \leq \pi/2\sqrt{2}$. Recall that $r_0 = \log \cosh t(r_0)$. By the definition (1) of $d(r_0)$, and the fact that $t(r_0)$ is increasing and $\frac{\pi}{2} - r_0$ is decreasing in r_0 , it suffices to display some value of r_0 in $(0, \pi/4)$ at which both $t(r_0)$ and $\frac{\pi}{2} - r_0$ are less than $\pi/2\sqrt{2}$. In particular, $r_0 = \pi/6.5$ will do. \square

3. Reduction to model spaces

3.1. The construction

For a given r_0 , consider a point p_- in M_- at distance r_0 from the boundary B_- (see Fig. 2). Say \bar{s} is *allowable* for r_0 if a point lying at length \bar{s} along B_- from the foot is transversely visible from p_- . Let $r_- = r_-(\bar{s}, r_0)$ be the corresponding radial distance from p_- ; then the allowability condition is $r_- < t(r_0)$. For a given $\bar{s} > 0$, define $v = v(\bar{s})$ and $w = w(\bar{s})$ by the following two conditions (C_{\pm}) (see Fig. 1):

(C_-) In M_- , v is the visual tangency distance to B_- from a point q_- at distance w from B_- ; that is, $v = t(w)$.

(C_+) In M_+ , v is the distance from a point q_+ satisfying $d(q_+, B_+) = w$ to a point lying at length \bar{s} along B_+ from the foot.

It can be verified (see Sect. 4) that v and w are increasing functions of \bar{s} for $\bar{s} < 1$.



Fig. 1.

Now fix $r_0 < .075$ and an allowable \bar{s} , and hence r_-, v and w . For r_0 in this range, $d(r_0) = t(r_0)$. As in the preceding section, consider a point p at distance r_0 from B , and set $M^* = M^*(p, r_0)$. Let γ denote a unitspeed B^* -geodesic that starts at a foot of p on B^* (so $\varphi = \pi/2$ at $s = 0$). Since \bar{s} is allowable for r_0 , it follows from the inequality $r_- \leq t(r_0)$ and the left sides of Lemma 2.2 (i) and (ii) that γ can be extended in B^* to length \bar{s} .

On the minimizer from p to the foot, take the point q at distance w from the foot (see Fig. 2). If we observe γ from q , then by (C_-) and the left side of Lemma 2.2 (i), γ remains visible either to its full length \bar{s} , or to a point m at radial distance v from q , whichever comes first. By (C_+) and the right sides of Lemma 2.2 (i) and (ii), the point m comes first, that is, the length s of γ out to m satisfies $s \leq \bar{s}$. Finally, if we again observe γ from p , since $s \leq \bar{s}$ and the radial distance from p increases along γ , the radial distance r from p to m is at most r_- . Thus in M^* we have a geodesic triangle $\triangle qpm$, lying entirely in the interior of M^* except for the vertex m , with two sidelengths $r_0 - w$ and v , and the third sidelength $r \leq r_-$.

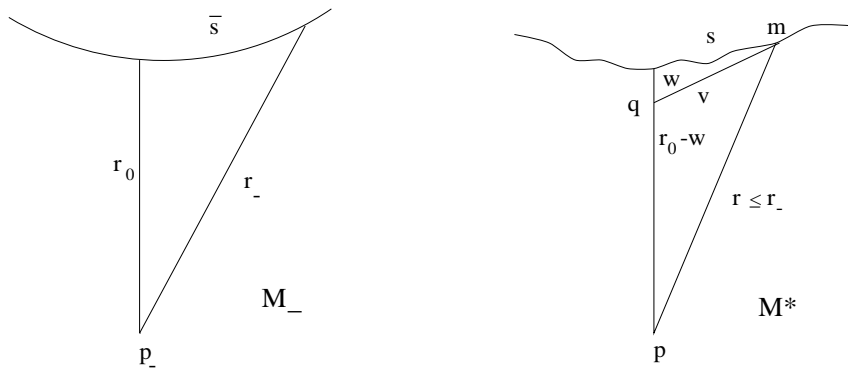


Fig. 2.

We shall prove in the next section:

Lemma 3.1. *Given $r_0 < .075$, there is an allowable \bar{s} such that for the triangle in S_{-1} with sidelengths $r_0 - w$, r_- and v , the angle opposite the side of length v is greater than $\pi/3$.*

Using this lemma, we can prove Theorems 1.1 and 1.2:

Proof of Theorem 1.1. First assume $\max\{\sup|K|, \sup\kappa^2\} \neq 0$, and rescale M so this maximum is 1. Suppose the inradius of M is less than .075. If p is a cut point of B in M , then $r_0 = d(p, B) < .075$. Let $M^* = M^*(p, r_0)$. Choose \bar{s} as in Lemma 3.1. Since \bar{s} is allowable for r_0 , we have just seen that B^* contains, from each footpoint x of p on B^* , a B^* -geodesic of length \bar{s} in every direction.

We claim that these geodesics sweep out a normal coordinate neighborhood of x in B^* . Otherwise, by a standard argument, there would be a nontrivial geodesic loop in B^* based at x with length at most $2\bar{s}$. Here, $\bar{s} \leq (1 - e^{-2r_0})^{1/2} \leq (1 - e^{-.15})^{1/2} \sim .37$, as may be verified by setting $r_- = t(r_0)$ in (6) below. But by hypothesis, the normal curvature in M^* of this loop is at most 1. Therefore by Lemma 2.4 (ii), according to which \bar{M}^* is a CAT(2)-space, and Theorem 2.3 (iii), the length of this loop must be at least the length of a circle of curvature 1 in S_2 , namely $\sqrt{2}\pi \sin(\tan^{-1} \sqrt{2}) = 2\pi/\sqrt{3}$, a contradiction.

Note that the closed normal coordinate balls of radius \bar{s} in B^* centered at two different footpoints x_1 and x_2 of p cannot intersect. If they did, then the sum of the two distances in B^* from x_1 and x_2 would have a minimum value on that intersection of closed balls. At such a minimum the two geodesics must join smoothly, forming a B^* -geodesic joining x_1 and x_2 . Otherwise we could move in the direction on B^* which bisects the angle between the two segments, thus shortening both while remaining in the intersection of the two balls. This contradicts Lemma 2.2, according to which the radial distance from p to a B^* -geodesic from a footpoint of p always increases.

Now, following our construction, consider a geodesic triangle $\triangle qpm$ in M^* , where q lies on a minimizer from p to B^* at distance $r_0 - w$ from p , m lies on a B^* -geodesic from the footpoint with $|qm| = v$, and $|pm| = r \leq r_-$ (see Fig. 2). We claim that the angle $\angle qpm$ in M^* is no less than its comparison angle in S_{-1} .

Since \bar{M}^* is a CAT(2)-space, its minimizers of length $< \pi/2\sqrt{2}$ are unique and vary continuously with their endpoints. However, these \bar{M}^* -geodesics may bifurcate if they touch B^* , signifying negatively infinite curvature. We consider the family of \bar{M}^* -minimizers $\{\beta_u\}$ from q to the points of the geodesic $\alpha = mp$, where $u \in [0, a]$ is the arclength parameter on α and $|\beta_u| < r_0 + t(r_0) < \pi/2\sqrt{2}$. In order to apply Toponogov's theorem, we need to know that this family does not touch B^* (except for the endpoint $\alpha(0) = m$). This will follow by applying the first variation formula and the convexity of radial distance, first from p , then from q , then from p again.

Specifically, since by Lemma 2.2 the radial distance from p to γ is strictly increasing along γ , the distance from p to β_0 is larger at the right endpoint than at the left. By convexity and first variation, the angle at $\alpha(0)$ between the velocity vectors of β_0 and α is obtuse. By Theorem 2.3 (ii) applied to the metric ball in \bar{M}^* with center q and radius $\pi/2\sqrt{2}$, the function $|\beta_u|$ is also convex in u . Furthermore, since β_u is varying uniquely and continuously in a geodesic metric space of curvature at most 2 and $|\beta_u| < \pi/2\sqrt{2}$, the first variation formula applies here also (see discussion in [AB]). Therefore the length $|\beta_u|$ decreases for $0 < u < c \leq a$, where either $c = a$ or the angle between the velocity vectors of

β_c and α is $\pi/2$. Since q lies at distance w from B^* and $|\beta_0| = v = t(w)$, Lemma 2.1 implies that for $0 < u \leq c$, the geodesic β_u does not touch the boundary B^* . For $u \geq c$, the angle between the velocity vectors of β_u and α cannot be obtuse. It follows that if we view these geodesics β_u from p , the convex radial distance function from p to β_u is nonincreasing. Therefore β_u lies entirely within radial distance $r_0 - w$ of p . Since r_0 is the distance from p to B^* , for $u \geq c$ the geodesic β_u also cannot touch B^* .

Knowing that the family $\{\beta_u\}$ joining q to $\alpha = mp$ runs through the interior of M^* , we can use an Alexandrov-Toponogov construction to prove $\angle qpm > \pi/3$. Consider the *Alexandrov development* in S_{-1} of α from q . By definition, this is a curve $\tilde{\alpha}$ in S_{-1} obtained as follows: associate to q a point \tilde{q} in S_{-1} , and associate to β_u a geodesic $\tilde{\beta}_u$ of the same length from \tilde{q} to $\tilde{\alpha}(u)$, turning monotonically in u so that u is also the arclength parameter of $\tilde{\alpha}$. Such a development is always possible, and unique up to a motion. Furthermore, since the curvature of the interior of M^* is no less than -1 , the development $\tilde{\alpha}$ is *convex* in the sense that for sufficiently small ϵ , the curvilinear triangle bounded by $\tilde{\beta}_{u \pm \epsilon}$ and $\tilde{\alpha}[u - \epsilon, u + \epsilon]$ is convex. (See [PP, §1.4] for a further discussion.)

Note that by first variation, the angle $\angle qpm$ at p between the velocity vectors of β_a and α is equal to the angle at $\tilde{\alpha}(a)$ between the velocity vectors of $\tilde{\beta}_a$ and $\tilde{\alpha}$. Now take the evolute of the convex curve $\tilde{\alpha}$ from its endpoint $\tilde{\alpha}(a)$. In this way we obtain a geodesic triangle in S_{-1} whose sidelengths adjacent to $\tilde{\alpha}(a)$ are $r_0 - w$ and $|\tilde{\alpha}| = r$ respectively, whose angle at $\tilde{\alpha}(a)$ is equal to $\angle qpm$, and whose opposite sidelength is at least $|\beta_0| = v$. By the hinge theorem, the triangle in S_{-1} of sidelengths $r_0 - w$, r and v has angle opposite the side of length v no greater than $\angle qpm$. By Lemma 2.2, $r_0 - w < r_0 < r$. It follows that in the latter triangle, the angle opposite the side of length $r_0 - w$ is acute. But then since $r \leq r_-$, the triangle in S_{-1} of sidelengths $r_0 - w$, r_- and v also has angle opposite the side of length v no greater than $\angle qpm$. Therefore by Lemma 3.1, $\angle qpm$ exceeds $\pi/3$.

To summarize, centered at the footpoints x of p there are disjoint closed normal coordinate balls N^* in B^* of radius \bar{s} . Each B^* -geodesic γ of length \bar{s} from a footpoint x contains a subarc from x to the first point m on γ for which $|qm| = v$. Let N_0^* be the portion of B^* swept out by these subarcs. Applying Lemma 2.1, with q and w replacing p and r_0 , we find that N_0^* is visible from q , since $v = t(w) = d(w)$. Moreover by Lemma 2.2 and first variation, the restriction to N_0^* of the distance function d_q from q is regular except at x , and its level hypersurface $H_{v-\epsilon}$ defined by $d_q = v - \epsilon$ intersects every geodesic from x transversely. Therefore $H_{v-\epsilon}$ is an imbedded topological $(n-2)$ -sphere, separating the normal coordinate ball N^* .

It follows that the unit vectors at p pointing to $H_{v-\epsilon}$ separate the unit tangent sphere S_p . For ϵ sufficiently small, we have seen that the angle subtended at p by x and any point of the hypersurface $H_{v-\epsilon}$ is greater than $\pi/3$. We conclude that the closed $(n-2)$ -spheres of radius $\pi/3$ in S_p about vectors pointing to distinct footpoints of p are disjoint. Since for any three points on a unit sphere, there

exist two at distance at most $2\pi/3$ apart, p has at most two footpoints on B^* and hence on B .

Thus every cut point p of B has degree 2. Therefore the cut locus C of B is locally the zero set of a smooth function $F_1 - F_2$ where F_i is the distance to a connected neighborhood B_i in B . Since the gradients never coincide, C is a smooth hypersurface of M which is the image of B under the two-to-one local diffeomorphism mapping footpoints to cutpoints. A maximal connected subset B^o of B that maps into a component of C is a component of B covering a component of C . Thus there is either a homeomorphism or a double covering map from $B^o \times [-1, 1]$ onto M , which maps $B^o \times 0$ onto C and is a local diffeomorphism on $B^o \times [-1, 0]$ and $B^o \times [0, 1]$.

In the pullback structure, each vertical curve is a possibly nonsmooth join of gradient curves of the distance functions to $B^o \times \mp 1$ on $B^o \times [-1, 0]$ and $B^o \times [0, 1]$ respectively. Let V_1 and V_2 denote the corresponding gradient fields, each of which extends to a neighborhood U of $B^o \times 0$ in $B^o \times [-1, 1]$. By our previous work, we may choose U so that each of the two boundary components subtends a visual cone angle of at least $\pi/3$ at every point of U , and hence the angle between V_1 and V_2 at any $q \in U$ is at least $2\pi/3$. Let F be a function on $B^o \times [-1, 1]$ that is smooth in the pullback structure, whose values on $B^o \times [-1, 0]$ are in $[1/2, 1]$ and 1 off U , and whose values on $B^o \times [0, 1]$ are in $[0, 1/2]$ and 0 off U . Then $V = FV_1 - (1 - F)V_2$ determines a smooth vector field on $B^o \times [-1, 1]$ of bounded length and whose projections on the vertical directions are bounded away from 0. Therefore its integral curves run from one boundary component to the other. By normalizing their parametrization, we conclude that either M or its double cover is diffeomorphic to $B^o \times [-1, 1]$.

It only remains to check the ‘flat/flat’ case, where $\max\{\sup |K|, \sup \kappa^2\} = 0$ and $\text{inr} \in (0, \infty)$. Again let p be a cut point of the boundary B . In this case all of M is visible from p , and there is a flat strip M^* in $T_p M$ mapped locally isometrically onto M by the exponential map. Here, either M or its double cover is isometric to the product of B^o and an interval. \square

Proof of Theorem 1.2. The topological assumptions of the theorem ensure that the hypersurface H separates X into two components. Indeed, by Alexander duality [S, Sect. 6.9], $H_1(X, X - H; Z/2Z)$ is isomorphic to $H_c^{n-1}(H; Z/2Z)$, hence has dimension 1 since H is connected. Then the long exact sequence for the pair (X, H) shows that $H_0(X - H)$ has dimension 2.

The fact that H is imbedded properly and smoothly makes it clear that each component has boundary H . Now we apply arguments from the proof of Theorem 1.1 to the closure M of one of these components. If the inradius of M were less than .075, then the cut locus C of $B = H$ would be doubly covered by H under the cutpoint map since H is connected. Since every point of C lies at distance no more than the inradius from its preimages in H and H is properly imbedded in X , then C is also properly imbedded in X . Therefore C separates X into

two components, both of which intersect H . But then H is not connected, a contradiction. \square

4. Estimates

This section is devoted to the proof of Lemma 3.1. We begin with a qualitative version, from which it follows that Theorems 1.1 and 1.2 hold for some dimension-independent constant c . We finish with precise estimates on c . Let functions $w = w(\bar{s})$, $v = v(\bar{s})$ and $r_- = r_-(\bar{s}, r_0)$ be defined as in Sect. 3. The defining equations are:

$$\cosh r_- = ((1 + \bar{s}^2)e^{r_0} + e^{-r_0})/2 \quad (6)$$

$$e^w = \cosh v \quad (7)$$

$$\cos v = \cos^2(\bar{s}/\sqrt{2}) \cos w + \sin^2(\bar{s}/\sqrt{2}) \sin w. \quad (8)$$

The constraints are: $0 < w < r_0$, $\bar{s} > 0$ and the allowability condition $\cosh r_- < \cosh t(r_0) = e^{r_0}$.

For the S_{-1} -triangle with sidelengths $r_0 - w$, r_- and v , let $\theta = \theta(\bar{s}, r_0)$ denote the angle opposite the side of length v . By the hyperbolic law of cosines

$$\begin{aligned} \cos \theta &= \frac{\cosh r_- \cosh(r_0 - w) - \cosh v}{\sinh r_- \sinh(r_0 - w)} \\ &= \frac{(1 + \bar{s}^2)e^{4r_0} + e^{2r_0} - ((3 - \bar{s}^2)e^{2r_0} - 1)e^{2w}}{((1 + \bar{s}^2)^2 e^{4r_0} - 2(1 - \bar{s}^2)e^{2r_0} + 1)^{\frac{1}{2}}(e^{2r_0} - e^{2w})}. \end{aligned} \quad (9)$$

Viewing (8) as a power series equation in v^2, \bar{s}^2 and w , and expressing w as a power series in v^2 via (7), it is easy to see that v^2 and w are analytic in \bar{s}^2 . Specific calculation gives $w = \bar{s}^2/2 + O(\bar{s}^4)$, $e^{2w} = 1 + \bar{s}^2 + O(\bar{s}^4)$.

From (1) we can deduce the following qualitative version of Lemma 3.1: *for any $\epsilon > 0$ there is a $\delta > 0$ such that if $r_0 < \delta$, then $\theta(\bar{s}, r_0) > \pi/2 - \epsilon$ for some allowable \bar{s} .* The argument becomes clear if we “blow up” the singularity of $\theta(\bar{s}, r_0)$ at the origin by changing to variables (\bar{s}, a) where $e^{2r_0} = 1 + a\bar{s}$. Then the lines $r_0 = \text{constant}$ are carried to hyperbolas. For \bar{s} sufficiently small, the region $w < r_0 < \delta$ in the open first (\bar{s}, r_0) -quadrant is carried diffeomorphically onto a region in the open first (\bar{s}, a) -quadrant bounded above by the hyperbola $a\bar{s} = e^{2\delta} - 1$ and below by a curve (namely the image of $r_0 = w$) which lies below some line $a = C\bar{s}$. It is easily checked that imposing the allowability condition does not change this general description of the image region. Setting

$$(e^{2w} - 1 - \bar{s}^2)/\bar{s}^4 = b$$

converts equation (1) into

$$\cos \theta = \frac{a^2 + (1 + a^2 - 2b)\bar{s}^2 + a(1 - 3b)\bar{s}^3 + b\bar{s}^4 + ab\bar{s}^5}{(4 + a^2 + 6a\bar{s} + (1 + 2a^2)\bar{s}^2 + 2a\bar{s}^3 + a^2\bar{s}^4)^{\frac{1}{2}}(a - \bar{s} - b\bar{s}^3)}. \quad (10)$$

Note that the righthand side of (10) extends across the positive a -axis to an analytic function $f(\bar{s}, a)$ with $f(0, a) = a/\sqrt{4 + a^2}$. By the implicit function theorem, since $\partial f/\partial a(0, a) \neq 0$, the equation $\cos \theta = \text{constant}$ defines a as an analytic function of \bar{s} in a neighborhood of any point on the positive a -axis. Now our claim is immediate from the fact that such a curve intersects all hyperbolas $a\bar{s} = k$ for k sufficiently small.

Lemma 3.1 states a specific estimate, namely that for each $r_0 < .075$, there is an allowable \bar{s} for which $\theta > \pi/3$. Computer experiments indicate that the constant $c = .075$ given here is the best that can be achieved by our construction. We will prove the lemma by showing that $\theta > \pi/3$ along the curve $e^{2r_0} = 1 + .845\bar{s}$, which runs from the origin in the (\bar{s}, r_0) -plane to $(.1915, .0752)$. It is easily verified that if $a > .2$ and $0 < \bar{s} < .2$, then the expression F for $\cos \theta$ given by (10) satisfies $\partial F/\partial b < 0$. Below we show

$$b \geq -.024 \text{ for } 0 < \bar{s} < .2. \quad (11)$$

Thus it suffices to prove $F^2 = P/Q < 1/4$ when $a = .845, b = -.024$ and $0 < \bar{s} < .192$. This in turn follows from calculations showing that the resulting polynomial $Q - 4P$ in \bar{s} is positive at $\bar{s} = 0$ and strictly decreasing to its first positive root at $\bar{s} > .192$.

The rest of the section is devoted to proving (11); the argument for this is rather technical. Assume $\bar{s} < .2$. In M_+ , let w_0 be the projection of a B_+ -arc of length \bar{s} on a radius of B_+ (see Fig. 1). Then $\tan w_0 = \tan^2(\bar{s}/\sqrt{2})$. From the series expansions for w in terms of \bar{s} (from (7) and (8)) and for w_0 , there is an initial interval on which w is increasing and $w < w_0$. On this interval, first variation gives $dv/d\bar{s} = (-dw/d\bar{s}) \cos \alpha + \cos \beta < 1$ where α and β are as in Fig. 1. Thus while $w < w_0$ holds, $v < \bar{s}$ and w increases. In addition, since $dw/d\bar{s} = (dv/d\bar{s}) \tanh v$ by (7), v also increases and $dw/d\bar{s} < (dv/d\bar{s}) \tanh \bar{s} < \bar{s}$, hence $w < \bar{s}^2/2$. But since $\tan^2 \rho > \tan \rho^2$ for $0 < \rho < 1$, we have $\tan w_0 > \tan(\bar{s}^2/2)$ and hence $w < w_0$ for $\bar{s} < .2$. Thus we have variables in the following range:

$$z = \bar{s}^2 \leq .04, \quad u = v^2 \leq .04, \quad w \leq .02. \quad (12)$$

By (8), (12) and alternating series estimates, one can show that $\cos v$ is bounded above by $1 - z/2 + z^2/12 - .499 w^2 + .505 zw$ and below by $1 - u/2 + .0416 u^2$. Thus

$$z - z^2/6 - 1.01zw + .998w^2 \leq u - .082u^2. \quad (13)$$

Similarly from (7) one obtains both

$$u/2 + u^2/24 \leq w + .504w^2 \quad (14)$$

and $1 + w + w^2/2 \leq 1 + u/2 + 1.03u^2/24$, which together imply $2w \leq u$.

Now we convert (14) to a quadratic upper bound in w on u . Since $dv/d\bar{s} \neq 0$, (7) and (8) determine a curve in the (w, u) -plane, which lies under the hyperbola defined by the case of equality in (13). If $u = 2w + Aw^2$ is the parabola that makes second order contact with this hyperbola at the origin, then $A = .673$. Since distinct conics can have at most four intersections, counting multiplicities, and since this parabola lies above the hyperbola at $w = .2$, we can conclude that it does so on the entire w range. (If the parabola were under for $|w| < \epsilon$, then it would have two intersections in addition to a multiplicity three intersection at the origin; if it were under for $0 < w < \epsilon$ and over for $-\epsilon < w < 0$, then it would have one intersection in addition to the multiplicity four intersection at the origin.) Thus

$$2w \leq u \leq 2w + Aw^2. \quad (15)$$

Substituting (15) into (13) gives

$$z/2 - z^2/12 - .505zw \leq w - .325w^2. \quad (16)$$

By the same method as before, (16) may be converted to a quadratic lower bound in z on w :

$$z/2 + Bz^2 \leq w \quad (17)$$

where $B = -.255$. Now exponentiate the square of both sides of (17), and substitute the series expansion of the left side, truncated after the fourth order term. Replacing z^3 and z^4 by $.04z$ and $.0016z^2$ maintains the inequality and yields $1 + z - .024z^2 \leq e^{2w}$, which is (11).

5. Metric structure

As $\text{inr}^2 \max\{\sup |K|, \sup \kappa^2\} \rightarrow 0$, one can ask whether M or its double cover becomes arbitrarily Lipschitz close to a metric product (see [GLP] for a discussion of Lipschitz distance). This would force the ratio of supremum to infimum of distances from boundary points to their opposite boundary component to approach 1. A counterexample is provided by the region in an n -sphere between a great hypersphere σ and an eccentric small hypersphere that converges to σ in such a way that this ratio approaches ∞ . It is also easy to see that the stronger condition $\text{diam}^2 \max\{\sup |K|, \sup \kappa^2\} \rightarrow 0$ is not sufficient to force M close to a product structure.

Almost-metric splitting does occur uniformly locally, however, as we now show. As was proved in the previous section, given $\epsilon > 0$, there is a $\delta > 0$ such that if $r_0 < \delta$, then for some allowable \bar{s} , the triangle in S_{-1} with sidelengths $r_0 - w$, r_- and v has angle opposite the side of length v greater than $(\pi/2) - \epsilon$. Now suppose \tilde{M} satisfies $\text{inr}^2 \max\{\sup |K|, \sup \kappa^2\} < (\delta/2)^2$ for small δ . By Theorem 1.1, \tilde{M} is diffeomorphic to a product, where \tilde{M} denotes M or its double cover according to whether M has one or two boundary components respectively. Moreover, as in the proof of Theorem 1.1, if V_i are the gradient

fields of the distance functions from the two boundary components B_i of \tilde{M} , then $\angle(V_1, -V_2) < 2\epsilon$.

Normalize so that $\text{inr} = 1$. Then the supremum a of distances from points of B_2 to B_1 satisfies $a \leq 2$. Since the normal curvatures of the level hypersurfaces of the distance from B_i are controlled by 1-dimensional Riccati inequalities in the sectional curvature bounds (see discussion in [K, Sect. 1]), these normal curvatures are arbitrarily close to 0 for δ sufficiently small. Since the normal curvatures are the logarithmic derivatives of the lengths of geodesic projections between level hypersurfaces, the interior of \tilde{M} maps diffeomorphically with local distortion arbitrarily close to 1 to an open submanifold P of the metric product $B_1 \times [0, a]$, where

$$P = \{(p, t) : 0 < t < f(p)\}$$

and $f(p)$ is the length of the \tilde{M} -geodesic normal to B_1 at p . Therefore \tilde{M} is arbitrarily Lipschitz close to the closure of P in $B_1 \times [0, a]$. Furthermore, since the geodesics normal to B_1 strike B_2 at angles arbitrarily close to $\pi/2$, by first variation $|df|$ is arbitrarily small.

Thus we obtain the following corollary, where the notation is as in Theorem 1.1. Recall that for δ sufficiently small, \tilde{M} is diffeomorphic to a product of a manifold without boundary and an interval, where \tilde{M} denotes M or its double cover according to whether M has one or two boundary components.

Corollary 5.1. *Given $C > 0$ and $\epsilon > 0$, we may choose $\delta > 0$ (independently of dimension) such that if $\text{inr}^2 \max\{\sup |K|, \sup \kappa^2\} < \delta$, then any metric ball in \tilde{M} of radius $C \text{ inr}$ lies in an open submanifold with boundary in \tilde{M} that is within Lipschitz distance ϵ of a metric product of a manifold without boundary and a closed interval.*

References

- [ABB] S.B. Alexander, I.D. Berg, R.L. Bishop, Geometric curvature bounds in Riemannian manifolds with boundary *Trans. Amer. Math. Soc.* **339** (1993), 703–716
- [AB] S.B. Alexander, R.L. Bishop, Comparison theorems for curves of bounded geodesic curvature in metric spaces of curvature bounded above. *Differential Geometry and Its Applications* **6** (1996), 67–86
- [Av1] A.D. Alexandrov, A theorem on triangles in a metric space and some of its applications. *Trudy Mat. Inst. Steklov* **38** (1951), 5–23 (Russian)
- [Av2] A.D. Alexandrov, Über eine Verallgemeinerung der Riemannschen Geometrie. *Schr. Forschungsinst. Math.* **1** (1957), 33–84
- [Av3] A.D. Alexandrov, Ruled surfaces in metric spaces. *Vestnik Leningrad Univ.* **12** (1957), 5–26 (Russian)
- [Bd] C. Bavard, Le rayon d’injectivité des surfaces à courbure majorée. *J. Differential Geometry* **20** (1984), 137–154
- [Bo] Yu. D. Burago, The injectivity radius of surfaces whose curvature is bounded above. *Ukrain. Geometr. Sb.* **21** (1978), 10–14 (Russian)
- [BK] P. Buser, H. Karcher, Gromov’s Almost Flat Manifolds. *Soc. Math. de France, Astérisque Fascicules*, 1981
- [G] M. Gromov, Synthetic geometry in Riemannian manifolds. In: O. Lehto (ed.) *Proceedings of the International Congress of Mathematicians, Helsinki, 1978*, vol. 1, Adademia Scientiarum Fennica, 1980, pp. 415–420

- [GLP] M. Gromov, Structures Métriques pour les Variétés Riemanniennes, rédigé par J. Lafontaine, P. Pansu, CEDIC/Fernand, Paris 1981
- [HT] R. Howard, A. Treibergs, A reverse isoperimetric inequality, stability and extremal theorems for plane curves with bounded curvature. *Rocky Mountain J. Math.* **25** (1995), 635–684
- [J] J. Jost, Eine geometrische Bemerkung zu Sätzen über harmonische Abbildungen, die ein Dirichletproblem lösen. *Manus. Math.* **32** (1980), 51–57
- [K] H. Karcher, Riemannian comparison constructions. In: S.S. Chern (ed.) *Global Differential Geometry*, vol. 27, MAA Studies in Math., 1989
- [L1] V.N. Lagunov, On the largest ball included in a closed surface. *Siberian Math. J.* **1** (1960), 205–232 (Russian)
- [L2] V.N. Lagunov, A.I. Fet, On the largest ball included in a closed surface, II. *Siberian Math. J.* **2** (1961), 874–883 (Russian)
- [LF1] V.N. Lagunov, A.I. Fet, Extremal questions for surfaces of a given topological type. *Siberian Math. J.* **4** (1963), 145–176 (Russian)
- [LF2] V.N. Lagunov, A.I. Fet, Extremal questions for surfaces of a given topological type, II. *Siberian Math. J.* **6** (1965), 1026–1036 (Russian)
- [PP] G. Perelman, A. Petrunin, Quasigeodesics in multidimensional Alexandrov spaces. Preprint
- [PI] G. Pestov, V. Ionin, On the largest disk included in a closed curve *Doklady Ak. nauk SSSR* **127** (1959), 1170–1172 (Russian)
- [S] E.H. Spanier, *Algebraic Topology*. McGraw-Hill, New York, 1966