UC Davis
UC Davis Previously Published Works

Title

The Cartan—Hadamard conjecture and the Little Prince

Permalink

https://escholarship.org/uc/item/70d4c5bf

Journal

Revista Matemática Iberoamericana, 35(4)

ISSN

0213-2230

Authors

Kloeckner, Benoît R
Kuperberg, Greg

Publication Date

2019-07-08

DOI

10.4171/rmi/1082

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

arXiv:1303.3115v3 [math.DG] 13 Feb 2017

THE CARTAN-HADAMARD CONJECTURE AND THE LITTLE PRINCE

BENOÎT R. KLOECKNER AND GREG KUPERBERG

ABSTRACT. The generalized Cartan-Hadamard conjecture says that if Ω is a domain with fixed volume in a complete, simply connected Riemannian n-manifold M with sectional curvature $K \leqslant \kappa \leqslant 0$, then $\partial\Omega$ has the least possible boundary volume when Ω is a round n-ball with constant curvature $K = \kappa$. The case $n = 2$ and $\kappa = 0$ is an old result of Weil. We give a unified proof of this conjecture in dimensions $n = 2$ and $n = 4$ when $\kappa = 0$, and a special case of the conjecture for $\kappa < 0$ and a version for $\kappa > 0$. Our argument uses a new interpretation, based on optical transport, optimal transport, and linear programming, of Croke's proof for $n = 4$ and $\kappa = 0$. The generalization to $n = 4$ and $\kappa \neq 0$ is a new result. As Croke implicitly did, we relax the curvature condition $K \leqslant \kappa$ to a weaker candle condition $\mathrm{Candle}(\kappa)$ or $\mathrm{LCD}(\kappa)$.

We also find counterexamples to a naïve version of the Cartan-Hadamard conjecture: For every $\varepsilon > 0$, there is a Riemannian $\Omega \cong B^3$ with $(1-\varepsilon)$-pinched negative curvature, and with $|\partial\Omega|$ bounded by a function of ε and $|\Omega|$ arbitrarily large.

We begin with a pointwise isoperimetric problem called "the problem of the Little Prince." Its proof becomes part of the more general method.

CONTENTS

Supported by ANR grant "GMT" JCJC - SIMI 1 - ANR 2011 JS01 011 01.
Supported by NSF grant CCF #1013079 and CCF #1319245.

1. INTRODUCTION

In this article, we will prove new, sharp isoperimetric inequalities for a manifold with boundary Ω, or a domain in a manifold. Before turning to motivation and context, we state a special case of one of our main results (Theorem 1.4).

Theorem 1.1. *Let Ω be a compact Riemannian n-manifold with boundary, with $n \in \{2,4\}$. Suppose that Ω has unique geodesics, has sectional curvature bounded above by $+1$, and that the volume of Ω is at most half the volume of the sphere $\mathbb{S}^n$ of constant curvature* 1. *Then the volume of $\partial\Omega$ is at least the volume of the boundary of a spherical cap in $\mathbb{S}^n$ with the same volume as Ω.*

Here and in the sequel, we say that a manifold (possibly with boundary) has *unique geodesics* when every pair of points is connected by at most one Riemannian geodesic. (More precisely, at most one connecting curve γ with $\nabla_{\gamma'}\gamma' = 0$. We do not consider locally shortest curves that hug the boundary to be geodesics.)

1.1. The generalized Cartan-Hadamard conjecture.

An isoperimetric inequality has the form

$$|\partial\Omega| \geqslant I(|\Omega|) \tag{1}$$

where I is some function. (We use $|\cdot|$ to denote volume and $|\partial\cdot|$ to denote boundary volume or perimeter; see Section 2.1.) The largest function $I = I_M$ such that (1) holds for all domains of a Riemannian n-manifold M is called the *isoperimetric profile* of M.

Besides the intrinsic appeal of the isoperimetric profile and isoperimetric inequalities generally, they imply other important comparisons. For example, they yield estimates on the first eigenvalue $\lambda_1(\Omega)$ of the Laplace operator by the Faber-Krahn argument [Cha84]. As a second example, the first author has shown [Klo15] that they imply a lower bound on a certain isometric defect of a continuous map $\phi : M \to N$ between Riemannian manifolds. Both of these applications also yield sharp inequalities when the isoperimetric optimum is a metric ball, which will be the case for the main results in this article.

The isoperimetric profile is unknown for most manifolds. The main case in which it is known is when M is a complete, simply connected manifold with constant curvature. Let $X_{n,\kappa}$ be this manifold in dimension n with curvature κ, and let $I_{n,k}$ be its isoperimetric profile. In other words, $X_{n,\kappa} = \sqrt{\kappa}\mathbb{S}^n$ is a sphere of radius $\sqrt{\kappa}$ when $\kappa > 0$; $X_{n,0} = \mathbb{E}^n$ is a Euclidean space; and $X_{n,\kappa} = \sqrt{-\kappa}\mathbb{H}^n$ is a rescaled hyperbolic space when $\kappa < 0$. Then a metric ball $B_{n,\kappa}(r)$ has the least boundary volume among domains of a given volume. Thus the profile is given by

$$I_{n,\kappa}(|B_{n,\kappa}(r)|) = |\partial B_{n,\kappa}(r)|.$$

Moreover, the volume $|B_{n,\kappa}(r)|$ and its boundary volume $|\partial B_{n,\kappa}(r)|$ are easily computable.

Instead of calculating the isoperimetric profile of a given manifold, we can look for a sharp isoperimetric inequality in a class of manifolds. Since $I_{n,\kappa}(V)$ decreases as a function of κ for each fixed V, it is natural to consider manifolds whose sectional curvature bounded above by some κ. This motivates the following well-known conjecture.

Conjecture 1.2 (Generalized Cartan-Hadamard Conjecture)**.** *If M is a complete, simply connected n-manifold with sectional curvature K bounded above by some $\kappa \leqslant 0$, then every domain $\Omega \subseteq M$ satisfies*

$$|\partial\Omega| \geqslant I_{n,\kappa}(|\Omega|). \tag{2}$$

(If M is not simply connected, then there are many counterexamples. For example, we can let M be a closed, hyperbolic manifold and let $\Omega \subseteq M$ be the complement of a small ball.)

The history of Conjecture 1.2 is as follows [Oss78, Dru10, Ber03]. In 1926, Weil [Wei26] established Conjecture 1.2 when $n = 2$ and $\kappa = 0$ for Riemannian disks Ω, without assuming an ambient manifold M, thus answering a question of Paul Lévy. Weil's result was established independently by Beckenbach and Radó [BR33], who are sometimes credited with the result. When $n = 2$, the case of disks implies the result for other topologies of Ω in the presence of M. It was first established by Bol [Bol41] when $n = 2$ and $\kappa \neq 0$. Rather later, Conjecture 1.2 was mentioned by Aubin [Aub76] and Burago-Zalgaller for $\kappa \leqslant 0$ [BZ88], and by Gromov [Gro81, Gro99]. The case $\kappa = 0$ is called the Cartan-Hadamard conjecture, because a complete, simply connected manifold with $K \leqslant 0$ is called a Cartan-Hadamard manifold.

Soon afterward, Croke proved Conjecture 1.2 in dimension $n = 4$ with $\kappa = 0$ [Cro84]. Kleiner [Kle92] proved Conjecture 1.2 in dimension $n = 3$, for all $\kappa \leqslant 0$, by a completely different method. (See also Ritoré [Rit05].) Morgan and Johnson [MJ00] established Conjecture 1.2 for *small* domains (see also Druet [Dru02] where the curvature hypothesis is on *scalar* curvature); however their argument does not yield any explicit size condition.

Actually, Croke does not assume an ambient Cartan-Hadamard manifold M, only the more general hypothesis that Ω has unique geodesics. We believe that the hypotheses of Conjecture 1.2 are negotiable, and it has some generalization to $\kappa > 0$. But the conjecture is not as flexible as one might think; in particular, Conjecture 1.2 is false for Riemannian 3-balls. (See Theorem 1.9 below and Section 4.) With this in mind, we propose the following.

Conjecture 1.3. *If Ω is a manifold with boundary with unique geodesics, if its sectional curvature is bounded above by some $\kappa > 0$, and if $|\Omega| \leqslant |X_{n,\kappa}|/2$, then $|\partial\Omega| \geqslant I_{n,\kappa}(|\Omega|)$.*

The volume restriction in Conjecture 1.3 is justified for two reasons. First, the comparison ball in $X_{n,\kappa}$ only has unique geodesics when $|\Omega| < |X_{n,\kappa}|/2$. Second, Croke [Cro80] proved a curvature-free inequality, using only the condition of unique geodesics, that implies a sharp extension of Conjecture 1.3 when $|\Omega| \geqslant |X_{n,\kappa}|/2$ (Theorem 1.15).

Of course, one can extend Conjecture 1.3 to negative curvature bounds (and then the volume condition is vacuous). The resulting statement is strictly stronger than Conjecture 1.2, since every domain in a Cartan-Hadamard manifold has unique geodesics, but there are unique-geodesic manifolds that cannot embed in a Cartan-Hadamard manifold of the same dimension (Figure 3).

Another type of generalization of Conjecture 1.2 is one that assumes a bound on some other type of curvature. For example, Gromov [Gro81, Rem. 6.28$\frac{1}{2}$] suggests that Conjecture 1.2 still holds when $K \leqslant \kappa$ is replaced by

$$K \leqslant 0, \quad \mathrm{Ric} \leqslant (n-1)^2 \kappa g. \tag{3}$$

In fact, Gromov's formulation is ambiguous: He considers $\kappa = -1$ and writes Ricci $\leqslant -(n-1)$, which could mean either $\mathrm{Ric} \leqslant -(n-1)^2 g$ or $\mathrm{Ric} \leqslant -(n-1)g$. The latter inequality is false for complex hyperbolic spaces $\mathbb{C}\mathrm{H}^n$. The former is similar to our root-Ricci curvature condition; see below.

Meanwhile Croke [Cro84] only uses a non-local condition that we call Candle(0) rather than the curvature condition $K \leqslant 0$; we state this as Theorem 1.13.

Our previous work [KK15] subsumes both of these two generalizations. More precisely, most of our results will be stated in terms of two volume comparison conditions, Candle(κ) and LCD(κ); see Section 2.2 for their definitions. One can interpret our two main results below (in weakened form) without referring to Section 2.2 by replacing Candle(κ) and LCD(κ) by $K \leqslant \kappa$, since $K \leqslant \kappa \implies$ LCD(κ) is Günther's inequality [Gün60, BC64], while LCD(κ) $\implies$ Candle(κ) is elementary. When $\kappa \leqslant 0$, one can also replace the Candle(κ) and LCD(κ) hypotheses by the mixed curvature bounds (3). In [KK15] we introduced a general curvature bound on what we call the *root-Ricci curvature* $\sqrt{\mathrm{Ric}}$, which is more general than both $K \leqslant \kappa$ and (3), and we proved that this bound implies *LCD*(κ) and Candle(κ).

1.2. **Main results.** For simplicity, we will consider isoperimetric inequalities only for compact, smooth Riemannian manifolds Ω with smooth boundary $\partial\Omega$; or for compact, smooth domains Ω in Riemannian manifolds M. Our constructions will directly establish inequalities for all such Ω. We therefore don't have to assume a minimizer or prove that one exists. Our results automatically extend to any limit of smooth objects in a topology in which volume and boundary volume vary continuously, *e.g.*, to domains with piecewise smooth boundary. Note that our uniqueness result, Theorem 1.7, does not automatically generalize to a limit of smooth objects; but its proof might well generalize to some limits of this type.

Our two strongest results are in the next two subsections. They both include Croke's theorem in dimension $n = 4$ as a special case. Each theorem has a volume restriction that we can take to be vacuous when $\kappa = 0$.

1.2.1. *The positive case.*

Theorem 1.4. *Let Ω be a compact Riemannian manifold with boundary, of dimension $n \in \{2,4\}$. Suppose that Ω has unique geodesics and is* Candle(κ) *with $\kappa \geqslant 0$ (*e.g.*, $K \leqslant \kappa$), and that $|\Omega| \leqslant |X_{n,\kappa}|/2$. Then $|\partial\Omega| \geqslant I_{n,\kappa}(|\Omega|)$.*

This is our fully general version of Theorem 1.1. As mentioned, Theorem 1.15 provides an optimal extension of Theorem 1.4 to the case $|\Omega| \geqslant |X_{n,\kappa}|/2$. Observe that the volume condition is vacuous when $\kappa = 0$, so that Theorem 1.4 implies Croke's theorem 1.13.

1.2.2. *The negative case.* When κ is negative and $n = 4$, we only get a partial result. (But see Section 9.) To state it, we let $r_{n,\kappa}(V)$ be the radius of a ball of volume V in $X_{n,\kappa}$. We define $\mathrm{chord}(\Omega)$ to be the length of the longest geodesic in Ω; we have the elementary inequality

$$\mathrm{chord}(\Omega) \leqslant \mathrm{diam}(\Omega).$$

Theorem 1.5. *Let M be a Cartan-Hadamard manifold of dimension $n \in \{2,4\}$ which is* $\mathrm{LCD}(\kappa)$ *with $\kappa \leqslant 0$ (e.g., $K \leqslant \kappa$). Let Ω be a domain in M, and if $n = 4$, suppose that*

$$\tanh(\mathrm{chord}(\Omega)\sqrt{-\kappa})\tanh(r_{n,\kappa}(|\Omega|)\sqrt{-\kappa}) \leqslant \frac{1}{2}. \tag{4}$$

Then $|\partial\Omega| \geqslant I_{n,\kappa}(|\Omega|)$.

Actually, Theorem 1.5 only needs M to be convex with unique geodesics rather than Cartan-Hadamard. However, we do not know whether that is a more general hypothesis for Ω. (See Section 4.) Observe that (4) is vacuous when $\kappa = 0$, and thus Theorem 1.5 also implies Croke's theorem 1.13.

The smallness condition (4) means that Theorem 1.5 is only a partial solution to Conjecture 1.2 when $n = 4$. Note that since $\tanh(x) < 1$ for all x, it suffices that *either* the chord length *or* the volume of Ω is small. *I.e.*, it suffices that

$$\sqrt{-\kappa}\min\left(\mathrm{chord}(\Omega), r_{n,\kappa}(|\Omega|)\right) \leqslant \mathrm{arctanh}(\frac{1}{2}) = \frac{\log(3)}{2}.$$

If we think of Conjecture 1.2 as parametrized by dimension, volume, and the curvature bound κ, then Theorem 1.5 is a complete solution for a range of values of the parameters.

1.2.3. *Pointwise illumination.* We prove a pointwise inequality which, in dimension 2, generalizes Weil's isoperimetric inequality [Wei26]. We state it in terms of *illumination* of the boundary of a domain Ω by light sources lying in Ω, defined rigorously in Section 3.

Theorem 1.6. *Let Ω be a compact Riemannian n-manifold with boundary, with unique geodesics, and which is* $\mathrm{Candle}(0)$*; and let $p \in \partial\Omega$. If we fix the volume $|\Omega|$, then the illumination of p by a uniform light source in Ω is maximized when Ω is Euclidean and is given by the polar relation*

$$r \leqslant k\cos(\theta)^{1/(n-1)} \tag{5}$$

for some constant k, with p at the origin. In particular, in dimension $n = 2$, the optimum Ω is a round disk.

Theorem 1.6 generalizes the elementary Proposition 3.1, the problem of the Little Prince, which was part of the inspiration for the present work.

When $n = 2$, Theorem 1.6 shows that a Euclidean, round disk maximizes illumination simultaneously at all points of its boundary, and therefore maximizes the *average* illumination over the boundary. But, as a consequence of the divergence theorem, the *total* illumination over the boundary is proportional to $|\Omega|$. A Euclidean, round disk must therefore minimize $|\partial\Omega|$, which is precisely Weil's theorem.

1.2.4. *Equality cases.* We also characterize the equality cases in Theorems 1.4 and 1.5, with a moderate weakening when $\kappa = 0$.

Theorem 1.7. *Suppose that Ω is optimal in Theorem 1.4 or 1.5, again with $n \in \{2,4\}$. When $\kappa = 0$, suppose further that Ω is $\sqrt{\mathrm{Ric}}$ class 0. Then Ω is isometric to a metric ball in $X_{n,\kappa}$.*

Again, see Section 2.2 for the definition of root-Ricci curvature $\sqrt{\mathrm{Ric}}$. In particular, $\sqrt{\mathrm{Ric}}$ class 0 implies $\mathrm{Candle}(0)$, but it does not imply $K \leqslant 0$ when $n > 2$.

We will prove Theorem 1.7 in Section 8.1; see also Section 9.

1.2.5. *Relative inequalities and multiple images.* Choe [Cho03, Cho06] generalizes Weil's and Croke's theorems in dimensions 2 and 4 to a domain $\Omega \subseteq M$ which is outside of a convex domain C, which is allowed to share part of its boundary with ∂C; he then minimizes the boundary volume $|\partial\Omega \setminus \partial C|$. The optimum in both cases is half of a Euclidean ball.

Choe's method in dimension 4 is to consider reflecting geodesics that reflect from ∂C like light rays. (This dynamic is also called *billiards*, but we use optics as our principal metaphor.) Such an Ω cannot have unique reflecting geodesics; rather two points in Ω are connected by at most two geodesics. We generalize Choe's result by bounding the number of connecting geodesics by any positive integer.

Theorem 1.8. *Let Ω be a compact n-manifold with boundary with $n = 2$ or 4, let $\kappa \geqslant 0$, and let $W \subset \partial\Omega$ be a (possibly empty) $(n-1)$-dimensional submanifold. Suppose that Ω is* Candle(κ) *for geodesics that reflect from W as a mirror, and suppose that every pair of points in Ω can be linked by at most m (possibly reflecting) geodesics. Suppose also that*

$$|\Omega| \leqslant \frac{|X_{n,\kappa}|}{2m}.$$

Then

$$|\partial\Omega \setminus \partial W| \geqslant \frac{I_{n,\kappa}(m|\Omega|)}{m}. \tag{6}$$

Note that Günther's inequality generalizes to this case (Proposition 5.8): If Ω satisfies $K \leqslant \kappa$, and if the mirror region W is locally concave, then Ω is LCD(κ) and therefore Candle(κ) for reflecting geodesics.

Theorem 1.8 is sharp, as can be seen from various examples. Let G be a finite group that acts on the ball $B_{n,\kappa}(r)$ by isometries. Then the orbifold quotient $\Omega = B_{n,\kappa}(r)/G$ matches the bound of Theorem 1.8, if we take the reflection walls to be mirrors and if we take $m = |G|$. Although Ω has lower-dimensional strata where it fails to be a smooth manifold, we can remove thin neighborhoods of them and smooth all ridges to make a manifold with nearly the same volume and boundary volume.

We could state a version of Theorem 1.8 for $\kappa < 0$ using the LCD(κ) condition, but it would be much more restricted because we would require an ambient M in which every two points are connected by *exactly* m geodesics. We do not know any interesting example of such an M. (*E.g.*, if the boundary of M is totally geodesic, then this case is equivalent to simply doubling M and Ω across ∂M.)

1.2.6. *Counterexamples in dimension 3.* We find counterexamples to justify the hypotheses of an ambient Cartan-Hadamard manifold and unique geodesics in Conjectures 1.2 and 1.3. One might like to replace these geometric hypotheses by purely topological ones. However, we show that even if Ω is diffeomorphic to a ball, this does not imply any isoperimetric inequality.

Theorem 1.9. *For every $\varepsilon > 0$, there is a Riemannian 3-manifold $\Omega \cong B^3$ with $(1-\varepsilon)$-pinched negative curvature and with arbitrarily large volume $|\Omega|$ and bounded surface area $|\partial\Omega|$ (depending only on ε).*

Recall that a Riemannian manifold has *δ-pinched negative curvature* when its sectional curvature K satisfies $-1 \leqslant K \leqslant -\delta$ everywhere.

While the manifold Ω we construct in Theorem 1.9 has trivial topology, its geometry is decidedly non-trivial. Most of its volume consists of a truncated hyperbolic knot complement $S_3 \setminus J$ with constant curvature $K = \kappa \approx -1$. Such an Ω has closed geodesics, which strongly contradicts the property of unique geodesics. Informally, we can say that Ω is "a 3-ball that wants to be a hyperbolic knot complement".

Theorem 1.9 was inspired by Joel Hass' construction of a negatively curved ball with totally concave boundary [Has94]. Both constructions yield counterexamples to a Riemannian extension problem considered by Pigola and Veronelli [PV16]. In both cases, the ball Ω has closed geodesics; if Ω could extend to a complete manifold W that satisfies $K \leqslant -1+\varepsilon$ or even $K \leqslant 0$, then its univeral cover $M = \tilde{W}$ would be a Cartan-Hadamard manifold with closed geodesics, contradicting the Cartan-Hadamard theorem. The ultimate purpose of either construction also obstructs a Cartan-Hadamard extension. In Hass' case, a compact Ω in a Cartan-Hadamard manifold cannot have totally concave boundary. In our case, by Kleiner's isoperimetric inequality [Kle92], Ω cannot have arbitrarily large volume and bounded surface area.

It is not hard to convert the result of Theorem 1.9 to a complete refutation of any possible isoperimetric relation for negatively curved 3-balls.

Corollary 1.10. *For each $V, A > 0$, there is a Riemannian 3-ball Ω with $K \leqslant -1$ and with $|\Omega| = V$ and $|\partial\Omega| = A$.*

We sketch the proof of Corollary 1.10: Starting with $|\Omega| \gg V$ and $|\partial\Omega|$ bounded, we can rescale Ω to make $|\Omega| = V - \varepsilon$ and $|\partial\Omega| < A$. We can then increase $|\partial\Omega|$ while increasing $|\Omega|$ by an arbitrarily small constant by adding a long, thin finger to Ω. Pinched negative curvature is an interesting extra property. We do not know whether one can achieve $|\partial\Omega| \to 0$ with $|\Omega|$ bounded below, and with $(1-\varepsilon)$-pinched negative curvature.

1.3. **The linear programming model.** Our method to prove Theorems 1.4 and 1.5 (and indirectly Theorem 1.6) is a reinterpretation and generalization of Croke's argument, based on optical transport, optimal transport, and linear programming.

We simplify our manifold Ω to a measure μ_Ω on the set of triples (ℓ, α, β), where ℓ is length of a complete geodesic $\gamma \subseteq \Omega$ and α and β are its boundary angles. Thus μ_Ω is always a measure on the set $\mathbb{R}_{\geqslant 0} \times [0, \pi/2)^2$, regardless of the geometry or even the dimension of Ω. We then establish a set of linear constraints on μ_Ω, by combining Theorem 5.3 (more precisely equations (23) and (24)) with Lemmas 5.4, 5.5, and 5.6. The result is the basic LP Problem 6.1 and an extension 7.2. The constraints of the model depend on the volume $V = |\Omega|$ and the boundary volume $A = |\partial\Omega|$, among other parameters.

Given such a linear programming model, we can ask for which pairs (V, A) the model is *feasible*; *i.e.*, does there exist a measure μ that satisfies the constraints? On the one hand, this is a vastly simpler problem than the original Conjecture 1.2, an optimization over all possible domains Ω. On the other hand, the isoperimetric problem, minimizing A for any fixed V, becomes an interesting question in its own right in the linear model.

Regarding the first point, finite linear programming is entirely algorithmic: It can be solved in practice, and provably in polynomial time in general. Our linear programming models are infinite, which is more complicated and should technically be called convex programming. Nonetheless, each model has the special structure of optimal transport problems, with finitely many extra parameters. Optimal transport is even nicer than general linear programming. All of our models are algorithmic in principle. In fact, our proofs of optimality in the two most difficult cases are computer-assisted using Sage [Sage].

Regarding the second point, our model is successful in two different ways: First, even though it is a relaxation, it sometimes yields a sharp isoperimetric inequality, *i.e.*, Theorems 1.4, 1.5, and 1.8. Second, our models subsume several previously published isoperimetric inequalities. We mention six significant ones. Note that the first four, Theorems 1.11-1.14, are special cases of Theorems 1.4, 1.5, and 1.8 as mentioned in Section 1.2. The other two results are separate, but they also hold in our linear programming models.

Theorem 1.11 (Variation of Weil [Wei26] and Bol [Bol41])**.** *Let Ω be a compact Riemannian surface with curvature $K \leqslant \kappa \geqslant 0$ with unique geodesics, and suppose that $\kappa|\Omega| \leqslant 2\pi$. Then for fixed area $|\Omega|$, the perimeter $|\partial\Omega|$ is minimized when $|\Omega|$ has constant curvature $K = \kappa$ and is a geodesic ball.*

Theorem 1.12 (Variation of Bol [Bol41])**.** *Suppose that $\Omega \subseteq M$ is a domain in a Cartan-Hadamard surface M that satisfies $K \leqslant \kappa \leqslant 0$. Then for fixed area $|\Omega|$, the perimeter $|\partial\Omega|$ is minimized when $|\Omega|$ has constant curvature $K = \kappa$ and is a geodesic ball.*

Theorem 1.13 (Croke [Cro84])**.** *If Ω is a compact 4-manifold with boundary, with unique geodesics, and which is* $\mathrm{Candle}(0)$*, then for each fixed volume $|\Omega|$, the boundary volume $|\partial\Omega|$ is minimized when Ω is a Euclidean geodesic ball.*

Theorem 1.14 (Choe [Cho03, Cho06])**.** *Let M be a Cartan-Hadamard manifold of dimension $n \in \{2,4\}$, and let $\Omega \subseteq M$ be a domain whose interior is disjoint from a convex domain $C \subseteq M$. Then*

$$|\partial\Omega \smallsetminus \partial C| \geqslant \frac{I_{n,0}(2|\Omega|)}{2}.$$

Theorem 1.15 (Croke [Cro80])**.** *If Ω is an n-manifold with boundary with unique geodesics, then $|\partial\Omega| \geqslant |\partial Y_{n,\rho}|$ where $Y_{n,\rho}$ is a hemisphere with constant curvature ρ and ρ is chosen so that $|\Omega| = |Y_{n,\rho}|$.*

Note that when $|\Omega| \geqslant |X_{n,\kappa}|/2$, we obtain $\rho \leqslant \kappa$, so that Croke's inequality extends Theorem 1.4, as promised. See the end of Section 8.5.2 for further remarks about this result.

Theorem 1.16 (Yau [Yau75])**.** *Let M be a Cartan-Hadamard n-manifold which is* $\mathrm{LCD}(\kappa)$ *with $\kappa < 0$. Then every domain $\Omega \subseteq M$ satisfies*

$$|\partial\Omega| \geqslant (n-1)\sqrt{-\kappa}|\Omega|.$$

Finally, we state the result that our models subsume all of these bounds.

Theorem 1.17. *Let μ be a measure on $\mathbb{R}^+ \times [0,\pi/2)^2$ that satisfies LP Problem 6.1, with formal dimension n, formal curvature bound κ, formal volume $V(\mu)$ (defined by (39)), and formal boundary volume $A(\mu)$ (defined by (35)). Then μ satisfies Theorem 1.4 and therefore 1.11. If μ satisfies LP Problem 7.2, then it satisfies Theorems 1.16 and 1.5, and therefore 1.12. If μ satisfies LP Problem 8.3, then it satisfies 1.15. If μ satisfies the LP model 8.1, then it satisfies Theorem 1.8 and therefore 1.14.*

We will prove some cases of Theorem 1.17 in the course of proving our other results; the remaining cases will be done in Section 8.5.

Our linear programming models are similar to the important Delsarte linear programming method in the theory of error-correcting codes and sphere packings [Del72, CS93, CE03]. Delsarte's original result was that many previously known bounds for error-correcting codes are subsumed by a linear programming model. But his model also implies new bounds, including sharp bounds. For example, consider the kissing number problem for a sphere in n Euclidean dimensions [CS93]. The geometric maximum is of course an integer, but in a linear programming model this may no longer be true. Nonetheless, Odlyzko and Sloane [OS79] established a sharp geometric bound in the Delsarte model, which happens to be an integer and the correct one, in dimensions 2, 8, and 24. (The bounds are, respectively, 6, 240, and 196,560 kissing spheres.) The basic Delsarte bound for the sphere kissing problem is quite strong in other dimensions, but it is not usually an integer and not usually sharp even if rounded down to an integer.

Another interesting common feature of the Delsarte method and ours is that they are both sets of linear constraints satisfied by a two-point correlation function, *i.e.*, a measure derived from taking pairs of points in the geometry.

1.4. **Other results.**

1.4.1. *Croke in all dimensions.* There is a natural version of Croke's theorem in all dimensions. This is a generalized, sharp isoperimetric inequality in which the volume of Ω is replaced by some other functional when the dimension $n \neq 4$. This result might not really be new; we state it here to further illustrate of our linear programming model.

If Ω is a manifold with boundary and unique geodesics, then the space G of geodesic chords in Ω carries a natural measure μ_G, called Liouville measure or étendue (Section 5).

Theorem 1.18. *Let Ω be a compact manifold with boundary of dimension $n \geqslant 4$, with unique geodesics, and with non-positive sectional curvature. Let*

$$L(\Omega) = \int_G \ell(\gamma)^{n-3}\,\mathrm{d}\mu_G(\gamma)$$

If $B_{n,0}(r)$ is the round, Euclidean ball such that

$$L(\Omega) = L(B_{n,0}(r)),$$

then

$$|\partial\Omega| \geqslant |\partial B_{n,0}(r)|.$$

By Theorem 5.3 (Santaló's equality),

$$\omega_{n-1}|\Omega| = \int_G \ell(\gamma)\,\mathrm{d}\mu_G(\gamma).$$

(Here $\omega_n = |X_{n,1}|$ is the n-sphere volume; see Section 2.1.) Thus Theorem 1.18 is Croke's Theorem if $n = 4$. The theorem is plainly a sharp isoperimetric bound for the boundary volume $|\partial\Omega|$ in all cases given the value of $L(\Omega)$, which happens to be proportional to the volume $|\Omega|$ only when $n = 4$.

Similar results are possible with a curvature bound $K < \kappa$, only with more complicated integrands $F(\ell)$ over the space G.

1.4.2. *Non-sharp bounds and future work.* We mention three cases in which the methods of this paper yield improved non-sharp results.

First, when $n = 3$ and $\kappa = 0$, Problem 6.1 yields a non-sharp version of Kleiner's theorem under the weaker hypotheses of Candle(0) and unique geodesics. Croke [Cro84] established the isoperimetric inequality in this case up to a factor of $\sqrt[3]{36/32} = 1.040\ldots$. Meanwhile Theorem 1.6 implies the same isoperimetric inequality up to a factor of $\sqrt[3]{27/25} = 1.026\ldots$. The wrinkle is that Croke's proof uses only (36), while Theorem 1.6 uses only (37). A combined linear programming problem should produce a superior if still non-sharp bound.

Second, it is a well-known conjecture that a metric ball is the unique optimum to the isoperimetric problem for domains in the complex hyperbolic plane $\mathbb{C}\mathrm{H}^2$. (The same conjecture is proposed for any non-positively curved symmetric space of rank 1.) If we normalize the metric on $\mathbb{C}\mathrm{H}^2$ so that it is $(-4,-1)$-pinched (Section 2.1), then $\mathbb{C}\mathrm{H}^2$ is

$\mathrm{LCD}(-16/9)$. Then Theorem 1.5 is, to our knowledge, better than what was previously established for moderately small volumes. Even so, this is a crude bound; we could do even better with a version of Problem 7.2 that uses the specific candle function of $\mathbb{C}\mathrm{H}^2$.

Third, even for domains in Cartan-Hadamard manifolds with $K \leqslant -1$ (or more generally $\mathrm{LCD}(-1)$), we can relax the smallness condition (4) in Theorem 1.5 simply by increasing the curvature bound κ from $\kappa = -1$. This is still a good bound for a range of volumes until it is eventually surpassed by Theorem 1.16. This is also a crude bound that can surely be improved, given that both Theorem 1.5 and Theorem 1.16 hold in the same linear programming model, Problem 7.2.

ACKNOWLEDGMENTS

The authors would like to thank Sylvain Gallot, Joel Hass, Misha Kapovich, and Qinglan Xia for useful discussions about Riemannian geometry and optimal transport. The authors would also like to thank an anonymous referee for detailed corrections and comments.

2. CONVENTIONS

2.1. **Basic conventions.** If $f : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ is an integrable function, we let

$$f^{(-1)}(x) \stackrel{\text{def}}{=} \int_0^x f(t)\,\mathrm{d}t$$

be its antiderivative that vanishes at 0, and then by induction its nth antiderivative $f^{(-n)}$. This is in keeping with the standard notation that $f^{(n)}$ is the nth derivative of f for $n > 0$.

If M is a Riemannian manifold, we let ν_M denote the Riemannian measure on M. As usual, TM is the tangent bundle of M, while we use UM to denote the unit tangent bundle. Also, if Ω is a manifold with boundary $\partial\Omega$, then we let

$$U^+\partial\Omega \stackrel{\text{def}}{=} \big\{u = (p,v) \mid p \in \partial\Omega, v \in U_p\Omega \text{ inward pointing}\big\}.$$

We say that M has (δ_1,δ_2)-*pinched curvature* if its sectional curvature K satisfies $\delta_1 \leqslant K \leqslant \delta_2$ everywhere. To paraphrase, we may say that M is pinched, its metric is pinched, etc.

We let $|M|$ be the volume of M:

$$|M| \stackrel{\text{def}}{=} \int_M \mathrm{d}\nu_M.$$

We let

$$\omega_n = |X_{n,1}| = \frac{2\pi^{(n+1)/2}}{\Gamma(\frac{n+1}{2})}$$

be the volume of the unit n-dimensional sphere $X_{n,1} = S^n \subseteq \mathbb{R}^{n+1}$.

2.2. **Candles.** Our main results are stated in terms of conditions $\mathrm{Candle}(\kappa)$ and $\mathrm{LCD}(\kappa)$ that follow from the sectional curvature condition $K \leqslant \kappa$ by Günther's comparison theorem [Gün60, BC64]. These conditions are non-local, but in previous work [KK15], we showed that they follow from another local condition, more general than $K \leqslant \kappa$ that we called $\sqrt{\mathrm{Ric}}$ class (ρ,κ). The original motivation is that Croke's theorem only needs that the manifold D is $\mathrm{Candle}(0)$, and even then only for pairs of boundary points. Informally, a Riemannian manifold M is $\mathrm{Candle}(\kappa)$ if a candle at any given distance r from an observer is dimmer than it would be at distance r in a geometry of constant curvature κ.

More rigorously, let M be a Riemannian manifold and let $\gamma = \gamma_u$ be a geodesic in M that begins at $p = \gamma(0)$ with initial velocity $u \in U_pM$. Then the *candle function* $j_M(\gamma, r)$ of M is by definition the normalized Jacobian of the exponential map

$$u \mapsto \gamma_u(r) = \exp_p(ru),$$

given by the equation

$$\mathrm{d}\nu_M(\gamma_u(r)) = j_M(\gamma_u, r)\,\mathrm{d}\nu_{U_pM}(u)\,\mathrm{d}r$$

for $r > 0$, where ν_M is the Riemannian volume on M and ν_{U_pM} is the Riemannian measure on the round unit sphere U_pM. More generally, if $a < b$, we define

$$j_M(\gamma, a, b) = j_M(\gamma_a, b-a),$$

where γ_a is the same geodesic as γ but with parameter shifted by a. We also define

$$j_M(\gamma, b, a) = j_M(\overline{\gamma_b}, b-a),$$

where $\overline{\gamma_b}$ is the same geodesic as γ, but reversed and based at $\gamma(b)$. (But see Corollary 5.2.)

The candle function of the constant-curvature geometry $X_{n,\kappa}$ is independent of the geodesic. We denote it by $s_{n,\kappa}(r)$; it is given by the following explicit formulas:

$$s_{n,\kappa}(r) = \begin{cases} \Big(\dfrac{\sin(r\sqrt{\kappa})}{\sqrt{\kappa}}\Big)^{n-1} & \text{if } \kappa > 0,\ r \leqslant \dfrac{\pi}{\sqrt{\kappa}} \\ r^{n-1} & \text{if } \kappa = 0 \\ \Big(\dfrac{\sinh(r\sqrt{-\kappa})}{\sqrt{-\kappa}}\Big)^{n-1} & \text{if } \kappa < 0. \end{cases} \tag{7}$$

We will also need the extension $s_{n,\kappa}(r) = 0$ when $\kappa > 0$ and $r \geqslant \pi/\sqrt{\kappa}$.

Definition. An n-manifold M is $\mathrm{Candle}(\kappa)$ if

$$j_M(\gamma, r) \geqslant s_{n,\kappa}(r)$$

for all γ and r. It is $\mathrm{LCD}(\kappa)$, for *logarithmic candle derivative*, if

$$\log(j_M(\gamma, r))' \geqslant \log(s_{n,\kappa}(r))'$$

for all γ and r. (Here the derivative is with respect to r.) The $\mathrm{LCD}(\kappa)$ condition implies the $\mathrm{Candle}(\kappa)$ condition by integration. If $\kappa > 0$, then these conditions are only required up to the focal distance $\pi/\sqrt{\kappa}$ in the comparison geometry.

To illustrate how $\mathrm{Candle}(\kappa)$ is more general than $K \leqslant \kappa$, we mention *root-Ricci curvature* [KK15]. Suppose that M is a manifold such that $K \leqslant 0$ and let $\kappa < 0$. For any unit tangent vector $u \in U_pM$ with $p \in M$, we define

$$\sqrt{\mathrm{Ric}}(u) \stackrel{\mathrm{def}}{=} \mathrm{Tr}(\sqrt{-R(\cdot, u, \cdot, u)}).$$

Here $R(u, v, w, x)$ is the Riemann curvature tensor expressed as a tetralinear form, and the square root is the positive square root of a positive semidefinite matrix or operator. We say that *M is of $\sqrt{\mathrm{Ric}}$ class κ* if $K \leqslant 0$ and

$$\sqrt{\mathrm{Ric}}(u) \geqslant (n-1)\sqrt{-\kappa}.$$

Then

$$K \leqslant \kappa \implies \sqrt{\mathrm{Ric}} \text{ class } \kappa \implies \mathrm{LCD}(\kappa) \implies \mathrm{Candle}(\kappa) \implies \mathrm{Ric} \leqslant (n-1)\kappa g.$$

The second implication, from $\sqrt{\mathrm{Ric}}$ to LCD, is the main result of [KK15]. (We also established a version of the result that applies for any $\kappa \in \mathbb{R}$. This version uses a generalized $\sqrt{\mathrm{Ric}}$ class (ρ, κ) condition that also requires $K \leqslant \rho$ for a constant $\rho > \max(\kappa, 0)$.) All implications are strict when $n > 2$. By contrast in dimension 2, the last condition trivially equals the first one, so all of the conditions are equivalent.

We conclude with two examples of 4-manifolds of $\sqrt{\mathrm{Ric}}$ class -1, and which are therefore LCD(-1), but that do not satisfy $K \leqslant -1$:

- The complex hyperbolic plane, normalized to be $(-\frac{9}{4}, -\frac{9}{16})$-pinched.
- The product of two simply connected surfaces that each satisfy $K < -9$.

Actually, the most important regime where Candle(κ) is weaker than $K \leqslant \kappa$ is at short distances. Since

$$j_M(\gamma, r) = r^{n-1} - \frac{\mathrm{Ric}(\gamma'(0), \gamma'(0))}{6} r^{n+1} + O(r^{n+2})$$

in dimension n, we can write informally that

$$\mathrm{Candle}(\kappa) \overset{\approx}{\iff} \mathrm{Ric} \leqslant (n-1)\kappa g$$

as $\mathrm{diam}(M) \to 0$.

3. The Little Prince and other stories

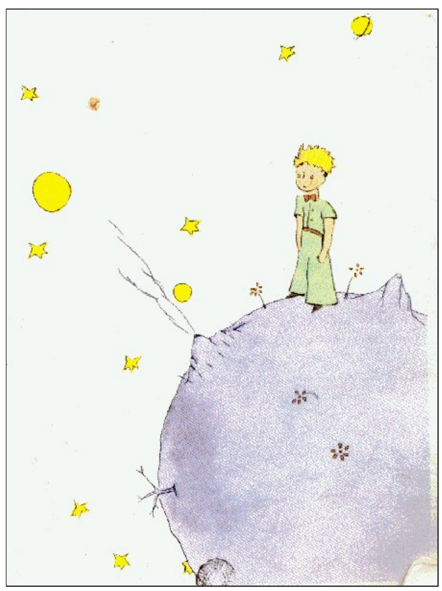

FIGURE 1. The Little Prince on his not-very-big planet, actually an asteroid.

3.1. **The problem of the Little Prince.** As Saint-Exupéry related to inhabitants of our planet, the Little Prince lives on his own planet, also known as asteroid B-612 (Figure 1). Since this planet is not very big, its gravitational pull is small and its habitation is precarious. The question arises as to what shape it should be to maximize the normal component of gravity for the Little Prince, assuming that the planet has a fixed mass, and a uniform, fixed mass density. Let Ω be the shape of the planet. The divergence theorem tells us that the average normal gravity is proportional to $|\Omega|/|\partial\Omega|$, so maximizing the average would be exactly the isoperimetric problem. Suppose instead that the Little Prince has a favorite location, and does not mind less gravity elsewhere. (After all, in the illustrations he usually stands on top of the planet.)

We cannot be sure of the dimension of the Little Prince or his planet. The illustrations are 2-dimensional, but the Prince visits the Sahara Desert which suggests that he is 3-dimensional. In any case higher-dimensional universes, which are a fashionable topic in physics these days, would each presumably have their own Little Prince. So we assume that the Little Prince is n-dimensional for some $n \geqslant 2$. We first assume Newtonian gravity and therefore a Euclidean planet; recall that in n dimensions, a divergenceless central gravitational force is proportional to r^{1-n}.

Proposition 3.1 (Little Prince Problem)**.** *Let Ω be the shape of a planet in n Euclidean dimensions with a pointwise gravitational force proportional to r^{1-n}. Suppose that the planet has a fixed volume $|\Omega|$ and a uniform, fixed mass density, and let $p \in \partial\Omega$. Then the total normal gravitational force $F(\Omega, p)$ at p is maximized when Ω is bounded by the surface $r = k\cos(\theta)^{1/(n-1)}$ for some constant k, in spherical coordinates centered at p.*

The problem of the Little Prince in 3 dimensions is sometimes used as an undergraduate physics exercise [McD03]. It has also been previously used to prove the isoperimetric inequality in 2 dimensions [HHM99]. However, our further goal is inequalities for curved spaces such as Theorem 1.6.

Proof. For convenience, we assume that the gravitational constant and the mass density of the planet are both 1. Given $x \in \Omega$, let $r = r(x)$ and $\theta = \theta(x)$ be the radius and first angle in spherical coordinates with the point p at the origin, and such that the normal component of gravity is in the direction $\theta = 0$. Then the total gravitational effect of a volume element $\mathrm{d}x$ at x is $\cos(\theta) r^{1-n}\,\mathrm{d}x$, so the total gravitational force is

$$F(\Omega, p) = \int_{x\in\Omega} \cos(\theta) r^{1-n}\,\mathrm{d}x.$$

In general, if $f(x)$ is a continuous function and we want to choose a region Ω with fixed volume to maximize

$$\int_\Omega f(x)\,\mathrm{d}x,$$

then by the "bathtub principle", Ω should be bounded by a level curve of f, *i.e.*,

$$\Omega = f^{-1}([k,\infty))$$

for some constant k. Our f is not continuous at the origin, but the principle still applies. Thus Ω is bounded by a surface of the form

$$r = k\cos(\theta)^{1/(n-1)}.$$ □

As explained above in words, the integral over $\partial\Omega$ of the normal component of gravity is proportional to $|\Omega|$ by the divergence theorem. More rigorously: We switch to a vector expression for gravitational force and we do not assume that $p = 0$. Then

$$F(\Omega, p) = \int_\Omega (x-p)|x-p|^{-n}\,\mathrm{d}x.$$

Since for each fixed $x \in \mathrm{Int}(\Omega)$, the vector field $p \mapsto (x-p)|x-p|^{-n}$ is divergenceless except at its singularity, we have

$$\int_{\partial\Omega} \langle -w(p), x-p\rangle |x-p|^{-n}\,\mathrm{d}p = \omega_{n-1}$$

where $w(p)$ is the outward unit normal vector at p. Thus

$$\int_{\partial\Omega} \langle -w(p), F(\Omega,p)\rangle \, \mathrm{d}p = \omega_{n-1}|\Omega|$$

by switching integrals. Then

$$\omega_{n-1}|\Omega| \leqslant |\partial\Omega| F_{\max}, \tag{8}$$

where $F_{\max}$ is the upper bound established by Proposition 3.1.

In particular, when $n=2$, the optimum Ω is the polar plot of $r = k\cos(\theta)$, which is a round circle. In this case

$$\langle -w(p), F(\Omega,p)\rangle = F_{\max}$$

at all points simultaneously. Thus when $n=2$, equation (8) is exactly the sharp isoperimetric inequality (2).

3.2. **Illumination and Theorem 1.6.** Proposition 3.1 is close to a special case of Theorem 1.6. To make it an actual special case, we slightly change its mathematics and its interpretation, but we will retain the sharp isoperimetric corollary using the divergence theorem. Instead of the shape of a planet, we suppose that Ω is the shape of a uniformly lit room, and we let $I(\Omega,p)$ be the total intensity of light at a point on the wall $p \in \partial\Omega$. More rigorously, if $\mathrm{Vis}(\Omega,p)$ is the subset of Ω which is visible from p (assuming that the walls are opaque, but allowing geodesics to be continued when they meet the boundary tangentially so that $\mathrm{Vis}(\Omega,p)$ is closed), then

$$I(\Omega,p) = \int_{\mathrm{Vis}(\Omega,p)} \langle -w(p), x-p\rangle |x-p|^{-n} dx.$$

We still have

$$\int_{\mathrm{Vis}(\partial\Omega,x)} \langle -w(p), x-p\rangle |x-p|^{-n} \, \mathrm{d}p = \omega_{n-1}$$

and we can still exchange integrals. Moreover,

$$I(\Omega,p) = F(\Omega,p)$$

when Ω is convex. Thus, this variation of Proposition 3.1 is also true and also implies (2).

We now consider the case when Ω is a curved Riemannian manifold, that is, Theorem 1.6. The proof is a simplified version of the proof of Theorems 1.4 and 1.5. Before giving the proof, we give a rigorous definition of illumination in the curved setting. (The definition agrees with the natural geometric assumption that light rays travel along geodesics.)

Let Ω be a compact Riemannian n-manifold with boundary and unique geodesics. We define a Riemannian analogue of $p \mapsto -(x-p)|x-p|^{-n}$, changing sign here to match the illumination interpretation. Namely, for each fixed $x \in \mathrm{Int}(\Omega)$, we define a tangent vector field v_x as follows. If $y \in \mathrm{Vis}(\Omega,x)$, then we let γ be the geodesic with $\gamma(0) = x$ and $\gamma(r) = y$, and then let

$$v_x(y) = \frac{\gamma'(r)}{j_\Omega(\gamma,r)}.$$

If $y \notin \mathrm{Vis}(\Omega,x)$, then we let $v_x(y) = 0$. The motivation, as above, is that this formula describes the radiation from a point source of light at x to the rest of Ω.

We claim that $\operatorname{div} v_x = \omega_{n-1}\delta_x$ in a distributional sense, where δ_x is the Dirac measure at x, so that we can then use v_x in the divergence theorem. It is routine to check that this holds at x itself and at any point y where v_x is continuous. The only delicate case is when $y \in \partial\,\mathrm{Vis}(\Omega,x) \smallsetminus \partial\Omega$. The vector field v_x is not continuous at these points; however, it is parallel to $\partial\,\mathrm{Vis}(\Omega,x)$ and thus does not have any singular divergence.

We fix a point $p \in \partial\Omega$ and again let $w(p)$ be the outward unit normal vector to $\partial\Omega$ at p. Then the illumination at p is defined by

$$I(\Omega, p) = \int_{\mathrm{Vis}(\Omega,p)} \langle w(p), v_x(p)\rangle \, \mathrm{d}\nu_\Omega(x).$$

Proof of Theorem 1.6. First, we express $I(\Omega, p)$ as an integral over $U = U_p^+\partial\Omega$, the unit inward tangent vectors at p. Given $u \in U$, let $\ell(u)$ be the length of the maximal geodesic segment defined by u and let $\alpha(u)$ be the angle of u with the inward normal $-w(p)$. Then, in polar coordinates we get

$$\begin{aligned} I(\Omega, p) &= \int_U \int_0^{\ell(u)} \cos(\alpha(u))\, \mathrm{d}t\, \mathrm{d}\nu_U(u) \\ &= \int_U \ell(u) \cos(\alpha(u))\, \mathrm{d}\nu_U(u). \end{aligned}$$

The first equality expresses the fact that the norm $||v_x(p)||$ is the reciprocal of the Jacobian of the exponential map from p. In other words, it is based on an optical symmetry principle (Corollary 5.2): If two identical candles are at x and p, then each one looks exactly as bright from the position of the other one.

Second, the Candle(0) hypothesis tells us that

$$|\Omega| \geqslant |\operatorname{Vis}(\Omega, p)| = \int_{\mathrm{Vis}(\Omega,p)} \mathrm{d}\nu_\Omega(x) \geqslant \int_U \int_0^{\ell(u)} t^{n-1}\, \mathrm{d}t\, \mathrm{d}\nu_U(u),$$

so that

$$|\Omega| \geqslant \int_U \frac{\ell(u)^n}{n}\, \mathrm{d}\nu_U(u). \tag{9}$$

Third, we apply the linear programming philosophy that will be important in the rest of the paper.

All of our integrands depend only on ℓ and α. Thus we can summarize all available information by projecting the measure ν_U to a measure

$$\sigma_\Omega = (\ell, \alpha)_*(\mathrm{d}\nu_U)$$

on the space of pairs

$$(\ell, \alpha) \in \mathbb{R}_{\geqslant 0} \times [0, \frac{\pi}{2}).$$

Then we want to maximize

$$I = \int_{\ell,\alpha} \ell \cos(\alpha)\, \mathrm{d}\sigma_\Omega \tag{10}$$

subject to the constraint

$$\int_{\ell,\alpha} \frac{\ell^n}{n}\, \mathrm{d}\sigma_\Omega \leqslant V. \tag{11}$$

We have one other linear piece of information: If we project volume on the hemisphere U into the angle coordinate $\alpha \in [0, \frac{\pi}{2})$, then the result is

$$\alpha_*(\mathrm{d}\sigma_\Omega) = \alpha_*(\mathrm{d}\nu_U) = \omega_{n-2} \sin(\alpha)^{n-2}\, \mathrm{d}\alpha, \tag{12}$$

since the latitude on U at angle α is an $(n-2)$-sphere with radius $\sin(\alpha)$.

We temporarily ignore geometry and maximize (10) for an abstract positive measure $\sigma = \sigma_\Omega$ that satisfies (11) and (12). To do this, choose $a > 0$, and let

$$f(\alpha) = \sup_{\ell > 0} \left(\ell \cos(\alpha) - \frac{a\ell^n}{n} \right). \tag{13}$$

We obtain

$$0 \leqslant \int_{\ell,\alpha} \left(f(\alpha) + \frac{a\ell^n}{n} - \ell\cos(\alpha) \right) \mathrm{d}\sigma(\ell,\alpha)$$
$$\leqslant \int_0^{\pi/2} f(\alpha)\omega_{n-2}\sin(\alpha)^{n-2}\,\mathrm{d}\alpha + aV - I. \tag{14}$$

The integral on the right side of (14) is a function of a only. Finally (14) is an upper bound on I, one that achieves equality if (11) is an equality and $\sigma = \sigma_\Omega$ is supported on the locus

$$\cos(\alpha) = a\ell^{n-1},$$

because that is the maximand of (13). The first condition tells us that Ω is Euclidean and visible from p. The second gives us the polar plot (5) if we take $k = a^{-1/(n-1)}$. □

Remark. It is illuminating to give an alternate Croke-style end to the proof of Theorem 1.6. Namely, Hölder's inequality says that

$$\begin{aligned} I &= \int_{\ell,\alpha} \ell\cos(\alpha)\,\mathrm{d}\sigma_\Omega \\ &\leqslant \left(\int_{\ell,\alpha} \ell^n\,\mathrm{d}\sigma_\Omega\right)^{\frac{1}{n}} \left(\int_{\ell,\alpha} \cos(\alpha)^{\frac{n}{n-1}}\,\mathrm{d}\sigma_\Omega\right)^{\frac{n-1}{n}} \\ &\leqslant (nV)^{\frac{1}{n}} \left(\int_0^{\frac{\pi}{2}} \cos(\alpha)^{\frac{n}{n-1}}\,\omega_{n-2}\sin(\alpha)^{n-2}\,\mathrm{d}\alpha\right)^{\frac{n-1}{n}}. \end{aligned}$$

The last expression depends only on V and n, while the inequality is an equality if (11) is an equality, and if

$$\ell^n \propto \cos(\alpha)^{\frac{n}{n-1}}.$$

The first condition again tells us that Ω is Euclidean and visible from p; the second one gives us the same promised shape (5).

The Croke-style argument looks simpler than our proof of Theorem 1.6, but what was elegance becomes misleading for our purposes. For one reason, our use of the auxiliary $f(\alpha)$ amounts to a proof of this special case of Hölder's inequality. Thus our argument is not really different; it is just another way to describe the linear optimization. For another, we will see more complicated linear programming problems in the full generality of Theorems 1.4 and 1.5 that do not reduce to Hölder's inequality.

4. Topology and geodesics

In this section we will analyze the effect of topology and geodesics on isoperimetric inequalities.

Weil and Bol established the sharp isoperimetric inequality (2) for Riemannian disks Ω with curvature $K \leqslant \kappa$, without assuming an ambient manifold M, and for any $\kappa \in \mathbb{R}$. The cases $\kappa \leqslant 0$ of the Weil and Bol theorems is equivalent to the $n = 2$ case of Conjecture 1.2 [Dru10].

The case $\kappa > 0$ is more delicate, even in 2 dimensions. Aubin [Aub76] assumed that B is a Riemannian ball with $K \leqslant \kappa$ and then that $\Omega \subseteq B$; but this formulation does not work. Even if B is a metric ball with an injective exponential map, and even if in addition $B \subset M$ and M is complete and simply connected with the same $K \leqslant \kappa$, there may be no control over the size of $\partial\Omega$. We can let M be a "barbell" consisting of two large, nearly round 2-spheres connected by a rod (Figure 2). Then Ω can be just the rod, while B is Ω union one end of the barbell. B is also a metric ball with an injective exponential map. Then Ω is an annulus $S^1 \times I$ in which both the meridian S^1 and the longitude I can have any length.

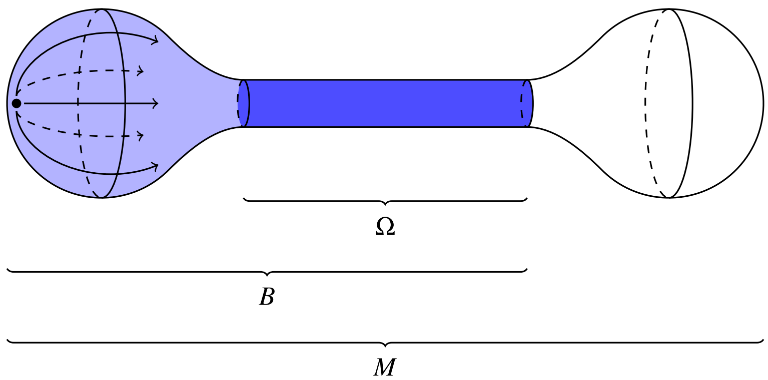

FIGURE 2. A counterexample $\Omega \subseteq B \subseteq M$ to Aubin's conjecture with $\kappa > 0$, in which both $|\Omega|$ and $|\partial\Omega|$ are unrestricted.

Thus both $|\Omega|$ and $|\partial\Omega|$ can have any value. Morgan and Johnson [MJ00] made the same point and used it to justify their small-volume hypothesis; of course, they require an upper bound on the volume that depends on the geometry of the ambient manifold.

Theorem 1.9 says that Weil's theorem fails completely for negatively curved Riemannian 3-balls. Our proof is similar to Hass's construction [Has94] of a negatively curved 3-ball with concave boundary.

If Ω is a smooth domain in a Cartan-Hadamard manifold as in Conjecture 1.2, then it has unique geodesics, but unique geodesics is a strictly weaker hypothesis even in 2 dimensions. For example, if Ω is a thin, locally Euclidean annulus with an angle deficit (Figure 3), then it has unique geodesics, but by the Gauss-Bonnet formula its inner circle cannot be filled by a non-positively-curved disk. Theorem 1.9 tells us that we need some geometric condition on a manifold Ω to obtain an isoperimetric inequality, because even the strictest topological condition, that Ω be diffeomorphic to a ball, is not enough. One natural condition is that Ω has unique geodesics. (See also Section 5.5 for a generalization.)

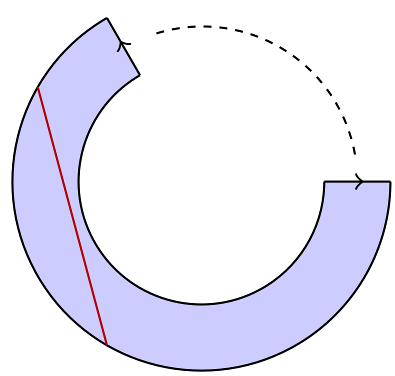

FIGURE 3. A conical, locally Euclidean annulus that has unique geodesics but does not embed in a Cartan-Hadamard surface, with a geodesic indicated in red. (We glue together the edges marked with arrows.)

Question 4.1. *If M is a Cartan-Hadamard manifold and Ω minimizes $|\partial\Omega|$ for some fixed value of $|\Omega|$, then is it convex? Is it a topological ball?*

Since we first proposed this question in an earlier version of the present article, Hass [Has16] proved that an isoperimetric minimizer Ω in a Cartan-Hadamard manifold need not be connected, which is thus a negative answer to both parts of the question. He also shows that in two dimensions, each connected component is a convex disk. He then gives partial evidence that nonconvex, connected minimizers exist in three dimensions. It may still be interesting to ask what restrictions there are on the topology of Ω, for simplicity given that it is a manifold with boundary.

In two dimensions, if Ω is a non-positively curved disk, then it has unique geodesics. (Proof: If a disk does not have unique geodesics, then it contains a geodesic "digon". By the Gauss-Bonnet theorem, a geodesic digon cannot have non-positive curvature.) Thus the $\kappa = 0$, $n = 2$ case of Theorem 1.4 implies Weil's theorem; but, as explained in the previous paragraph, it is more general.

In higher dimensions, there are non-positively curved smooth balls with closed geodesics. Hass's construction has closed geodesics, and so does our construction in Theorem 1.9.

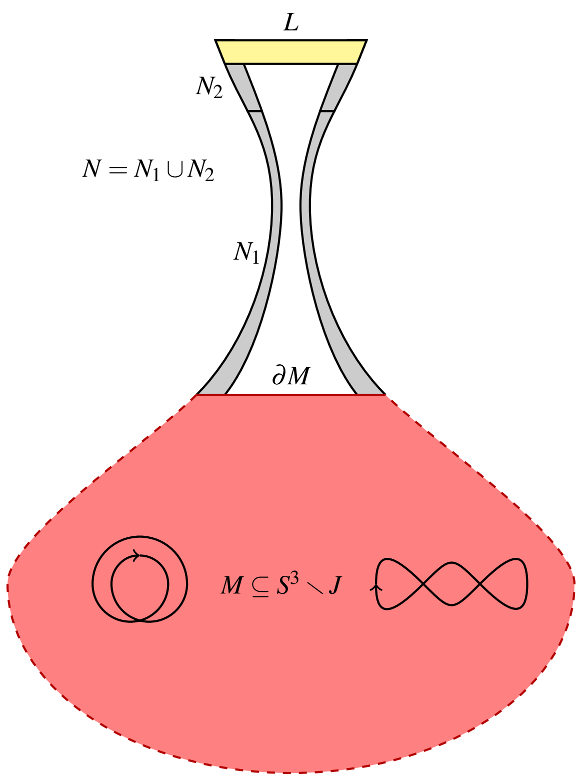

FIGURE 4. A diagram of $\Omega = M \cup N \cup L$, schematically like a decanter. It consists of a truncated hyperbolic knot complement M, plus a 2-handle $N \cup L$ that consists of a 3-ball lid L and a thickened annulus neck $N = N_1 \cup N_2$.

4.1. **Proof of Theorem 1.9.** In our proof of Theorem 1.9, we will construct Ω to be $(-1-\varepsilon,-1+\varepsilon)$-pinched, so that $\sqrt{1+\varepsilon}\Omega$ is $(-1,-1+2\varepsilon)$-pinched. We can then change ε to match the stated conclusion of Theorem 1.9.

Our construction is shown schematically in Figure 4. We make Ω as a union of three pieces M, N, and L. M is a truncated hyperbolic knot complement, L is a "lid" which is a horospheric pseudocylinder, and N is a connecting neck which is a thickened annulus. Both M and L have constant curvature $K=-1$, while the neck N has a $(-1-\varepsilon,-1+\varepsilon)$-pinched metric that interpolates between the metrics on ∂M and ∂L and meets each one along an annulus. Although the two ends of N are both horospheric annuli, they are mismatched in two ways: First, ∂M and the bottom of ∂L are both concave, so their extrinsic curvature must be interpolated. Second, the annulus $\partial M\cap N$ is *vertical* (isometric to a cylinder $S^1\times I$ with the product metric) while $\partial L\cap N$ is *horizontal* (isometric to an annulus in $\mathbb{E}^2$). In order to achieve both interpolations, we further divide $N=N_1\cup N_2$ into two thickened annuli N_1 and N_2.

To construct N_1 and N_2, which will be the most technical part of the proof, we review some facts about warped products [BO69, AB04]. Recall that if B and F are two Riemannian manifolds and $h:B\to\mathbb{R}_+$ is a smooth function, we can define a Riemannian metric on $B\times F$ by the formula

$$\mathrm{d}s^2_{B\times F}(p,q)=\mathrm{d}s^2_B(p)+f(p)^2\mathrm{d}s^2_F(q)$$

for $(p,q)\in B\times F$. The manifold $B\times F$ with this metric is denoted $B\times_f F$ and is called a *warped product*, while the function f is a *warping function*. In this paper, we will only need warped products of the form $I\times_f M$, where the base I is an interval.

Lemma 4.1. *Let I be an interval, let $f:I\to\mathbb{R}_+$ be a warping function, let F be a Riemannian manifold, and let $W=I\times_f F$. Then:*

1. *If F is locally Euclidean and $f(t)=e^t$, then W has constant curvature $K=-1$.*
2. *If F has constant curvature $K=-1$ and $f(t)=\cosh(t)$, then W also has constant curvature -1.*
3. *If F is 1-dimensional, then the intrinsic curvature of W is given by*

$$K_W(t,x)=-\frac{f''(t)}{f(t)}$$

for $t\in I$ and $x\in F$.

4. *If F is $(-1-\varepsilon,-1+\varepsilon)$-pinched and $f(t)=\cosh(t)$, then W is also $(-1-\varepsilon,-1+\varepsilon)$-pinched.*

In the proof of Lemma 4.1, and later in the proof of Theorem 1.9, we will make use of the standard upper half-space model for hyperbolic space:

$$\mathrm{d}s^2_{\mathbb{H}^{n+1}}=\frac{\mathrm{d}s^2_{\mathbb{E}^n}+\mathrm{d}z^2}{z^2}=\frac{\mathrm{d}x_1^2+\ldots+\mathrm{d}x_n^2+\mathrm{d}z^2}{z^2}. \tag{15}$$

Recall that in this model, $x_k\in\mathbb{R}$ for every k and $z>0$.

Proof. Cases 1, 2, and 3 all follow quickly from the conditions (16) and (17) below. However, as these are well-known facts in differential geometry, we also give separate calculations. Case 1 is confirmed by a standard metric model of $\mathbb{H}^{n+1}$,

$$\mathrm{d}s^2_{\mathbb{H}^{n+1}}=e^{2t}\mathrm{d}s^2_{\mathbb{E}^n}+\mathrm{d}t^2,$$

which is obtained from (15) by the change of variables $z = e^{-t}$. Case 2 is confirmed by a standard metric model of $\mathbb{H}^{n+1}$,

$$\mathrm{d}s^2 = \cosh(t)^2 \mathrm{d}s^2_{\mathbb{H}^n} + \mathrm{d}t^2,$$

which may also be obtained from (15) by the change of variables

$$(x_n, z) = (y\tanh(t), y\,\mathrm{sech}(t)),$$

to obtain the metric

$$\mathrm{d}s^2_{\mathbb{H}^{n+1}} = \cosh(t)^2 \frac{\mathrm{d}x_1^2 + \ldots + \mathrm{d}x_{n-1}^2 + \mathrm{d}y^2}{y^2} + \mathrm{d}t^2.$$

Case 3 follows from the Jacobi field equation (64), considering that the I fibers in a warped product $I \times_f F$ are geodesic curves.

Case 4 is a special case of a result of Alexander-Bishop [AB04, Prop 2.2.]. Reducing to the case of a one-dimensional base $B = I$, this proposition says that:

1. W has curvature bounded above by K_0 if

$$f'' \geqslant -K_0 f \qquad \text{and} \qquad K_F \leqslant K_0 f^2 + f'^2. \tag{16}$$

2. W has curvature bounded below by K_0 if

$$f'' \leqslant -K_0 f \qquad \text{and} \qquad K_F \geqslant K_0 f^2 + f'^2. \tag{17}$$

(The proposition states "if and only if" and requires that B be complete, but their proof makes clear that the "if" direction does not require completeness.) If we take $K_0 = -1+\varepsilon$ for the upper bound and $K_0 = -1-\varepsilon$ for the lower bound, and if $f(t) = \cosh(t)$, we obtain the requirements

$$\cosh(t)(1-\varepsilon) \leqslant \cosh(t) \leqslant \cosh(t)(1+\varepsilon)$$
$$-1-\varepsilon\cosh(t)^2 \leqslant K_F \leqslant -1+\varepsilon\cosh(t)^2.$$

All of these inequalities hold immediately. □

Lemma 4.2. *Let $\varepsilon, \delta, a, b > 0$. Then there exists $c > 2$ and a smooth $f : [0,c] \to \mathbb{R}_+$ such that the manifold*

$$N_1 = [-\delta, \delta] \times \mathbb{R}/2\pi\mathbb{Z} \times [0,c]$$

with coordinates (ρ, θ, h) and metric

$$\mathrm{d}s^2 = \mathrm{d}\rho^2 + \cosh(\rho)^2 (f(h)^2\,\mathrm{d}\theta^2 + \mathrm{d}h^2)$$

is $(-1-\varepsilon, -1+\varepsilon)$-pinched, and such that

$$f(h) = \begin{cases} ae^{-h} & h \leqslant 1 \\ be^{h-c} & h \geqslant c-1 \end{cases}$$

and

$$|\partial N_1| = O_\varepsilon\big(\cosh(\delta)^2(a+b)\big).$$

Proof. Following Lemma 4.1, N_1 is $(-1-\varepsilon, -1+\varepsilon)$-pinched if and only if

$$(1-\varepsilon)f(h) \leqslant f''(h) \leqslant (1+\varepsilon)f(h).$$

This relation holds if and only if the logarithmic derivative $u(h) = f'(h)/f(h)$ approximately satisfies a Riccati equation:

$$1-\varepsilon \leqslant u'(h) + u(h)^2 \leqslant 1+\varepsilon. \tag{18}$$

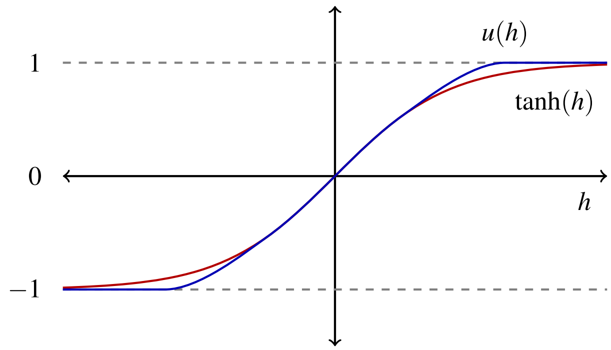

FIGURE 5. The function $u(h)$, an approximate Riccati solution that transitions from -1 to 1 and which partly agrees with the exact solution $\tanh(h)$

We first construct $u(h)$, for convenience for all $h \in \mathbb{R}$. Actually, we will shift h by a constant, which has no effect on (18). Let $u(h)$ be a smooth function such that:

1. $u(h) = \tanh(h)$ when $|\tanh(h)| \leqslant 1 - \frac{\varepsilon}{2}$.
2. $0 \leqslant u'(h) \leqslant \varepsilon$ when $|\tanh(h)| > 1 - \frac{\varepsilon}{2}$.
3. $u(h)$ increases with h until it reaches 1 at $h = c_0$ and then stays constant.
4. $u(-h) = -u(h)$ for all h.

(See Figure 5.) When $|\tanh(h)| \leqslant 1 - \frac{\varepsilon}{2}$, we have

$$u'(h) + u(h)^2 = 1.$$

For other values of h, we have

$$1 - \varepsilon < u(h)^2 \leqslant 1, \qquad 0 \leqslant u'(h) \leqslant \varepsilon,$$

so that (18) is satisfied in all cases.

Choose $c_1 \leqslant -c_0 - 1$ and $c_2 \geqslant c_0 + 1$ such that also

$$\int_{c_1}^{c_2} u(t)\,\mathrm{d}t = \log(b) - \log(a),$$

and let $c = c_2 - c_1$. Then

$$f(h) = a \exp\left(\int_{c_1}^{h+c_1} u(t)\,\mathrm{d}t\right)$$

has all of the required properties on the interval $[0,c]$.

We can estimate $|\partial N_1|$ by first considering the area of the level surface $\rho = 0$, which is $\int_0^c 2\pi f(h)\,\mathrm{d}h$. Splitting the integral of f into

$$\int_0^c f(h)\,\mathrm{d}h = \int_0^{-c_0-c_1} f(h)\,\mathrm{d}h + \int_{-c_0-c_1}^{c_0-c_1} f(h)\,\mathrm{d}h + \int_{c_0-c_1}^{c} f(h)\,\mathrm{d}h,$$

we observe that

$$\int_0^{-c_0-c_1} f(h)\,\mathrm{d}h = \int_0^{-c_0-c_1} ae^{-h}\,\mathrm{d}h < a$$
$$\int_{c_0-c_1}^{c} f(h)\,\mathrm{d}h = \int_{c_0-c_1}^{c} be^{h-c}\,\mathrm{d}h < b$$
$$\int_{-c_0-c_1}^{c_0-c_1} f(h)\,\mathrm{d}h < 2c_0 \max(a,b) = O_\varepsilon(a+b).$$

The estimate for $|\partial N_1|$ follows quickly. $\square$

Lemma 4.3. *Let $\varepsilon > 0$ and let $f : [0,3] \to [0,1]$ be a smooth function such that $f(h) = 0$ for $h \leqslant 1$ and $f(h) = 1$ for $h \geqslant 2$. Then there exists $b > 0$ such that the manifold*

$$N_2 = [-1,1] \times \mathbb{R}/2\pi\mathbb{Z} \times [0,3]$$

with coordinates (ρ, θ, h) and with the metric

$$\mathrm{d}s^2 = e^{2h}\big(\mathrm{d}\rho^2 + (b + f(h)\rho)^2\,\mathrm{d}\theta^2\big) + \mathrm{d}h^2, \tag{19}$$

is $(-1-\varepsilon, -1+\varepsilon)$-pinched.

Note that if $f(h)$ is locally constant near $h = h_0$, then Lemma 4.1 tells us that the metric (19) has constant negative curvature $K = -1$ at h_0. In particular, $K = -1$ for every $h \in [0,1] \cup [2,3]$.

Proof. For clarity, we work in the universal cover $\tilde{N}_2$, so that $\theta \in \mathbb{R}$. Without yet choosing b, we apply the change of variables $\theta = \alpha/b$ to obtain the metric

$$\mathrm{d}s^2 = e^{2h}\left(\mathrm{d}\rho^2 + \big(1 + \frac{f(h)\rho}{b}\big)^2\,\mathrm{d}\alpha^2\right) + \mathrm{d}h^2. \tag{20}$$

Since h and ρ are bounded independently of the choice of b, and since $f(h)$ is a fixed, smooth function, the metric (20) converges uniformly in the C^∞ topology to the metric

$$\mathrm{d}s^2 = e^{2h}(\mathrm{d}\rho^2 + \mathrm{d}\alpha^2) + \mathrm{d}h^2 \tag{21}$$

as $b \to \infty$. Recall that the curvature tensor R_{ijkl} of a manifold M with a metric g_{ij} has a polynomial formula in terms of the derivatives $g_{ij,k}$ and $g_{ij,kl}$, and the matrix inverse g^{ij}. In our case, the limiting metric (21) has constant curvature $K = -1$ by Lemma 4.1. Both g_{ij} and g^{ij} are uniformly bounded, since they are independent of the non-compact coordinate α. It follows that the sectional curvature of $\tilde{N}_2$ converges uniformly to $K = -1$ as $b \to \infty$. This is equivalent to the conclusion that $\tilde{N}_2$ or N_2 is $(-1-\varepsilon, -1+\varepsilon)$-pinched when b is sufficiently large. $\square$

Proof of Theorem 1.9. Let $J \subseteq S^3$ be a hyperbolic knot and give its complement $S^2 \smallsetminus J$ its complete hyperbolic metric with curvature $K = -1$. We can choose J so that $|S^3 \smallsetminus J|$ is arbitrarily high by a theorem of Adams [Ada05]. A tubular neighborhood of J is metrized as a parabolic cusp, and this cusp can be truncated to obtain a manifold M with a horospheric torus boundary ∂M. Let $\gamma \subset \partial M$ be a geodesic meridian circle. Note that we can truncate $|S^3 \smallsetminus J|$ as far out along the cusp as we like. We choose the truncation so that

$$|M| \geqslant |S^3 \smallsetminus J| - 1, \qquad |\partial M| \leqslant 1, \qquad |\gamma| \leqslant 1.$$

If we attach a 2-handle $D^2 \times I$ to M along γ, then the result

$$\Omega = M \cup (D^2 \times I)$$

is diffeomorphic to a ball B^3. (Recall that an n-dimensional k-handle is a $B^k \times B^{n-k}$ which is to be attached along $(\partial B^k) \times B^{n-k}$ to some other manifold.) We want to give the handle $D^2 \times I$ a $(-1-\varepsilon, -1+\varepsilon)$-pinched metric that extends smoothly to the metric on M. We also want to bound $|\partial(D^2 \times I)|$ by a constant, independent of the choice of M (but depending on ε).

We construct the 2-handle $D^2 \times I$ as the union of a thickened annulus N, the "neck"; and a 3-ball L, the "lid". The connecting neck N is divided into two stages, N_1 and N_2. Each N_k is of the form $I_{\text{short}} \times S^1 \times I_{\text{long}}$, where $S^1 \times I_{\text{long}}$ is thus a long annulus. More precisely

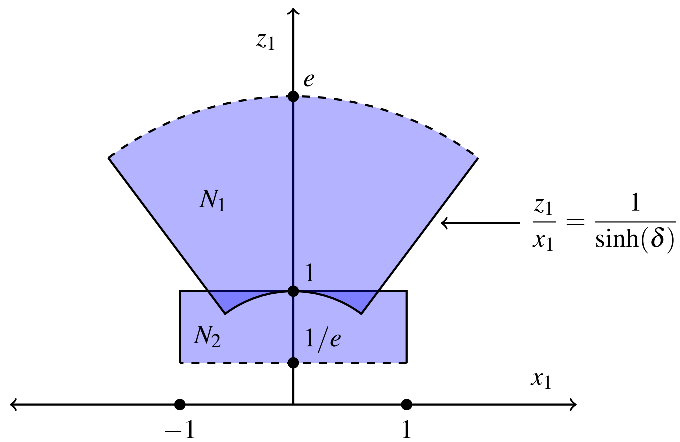

FIGURE 6. Gluing N_1 to N_2 in the half-space model in coordinates x_1 and z_2. If $\sinh(\delta) \leqslant 1$, then N_1 inserts into the end of N_2 as shown.

(as in Figure 4), we will glue M to N_1, N_1 to N_2, and N_2 to L. It will be convenient for each pair to overlap with positive volume.

We construct N_1 and N_2 using Lemmas 4.2 and 4.3, which both use the coordinates (ρ, θ, h). To distinguish them, we change the variables names to ρ_1 and h_1 in Lemma 4.2 and to ρ_2 and h_2 in Lemma 4.3; the coordinate θ will be the same.

We choose the constant a in Lemma 4.2 so that $|\gamma| = 2\pi a$, and we choose the constant b to match its value provided in Lemma 4.3. We parametrize γ by $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ so that $a\theta$ represents the length along γ from some starting point. Since γ is horocyclic, it has a neighborhood in M with coordinates (ρ_1, θ, h_1) with the metric

$$\mathrm{d}s^2 = \cosh(\rho_1)^2(a^2 e^{-2h_1}\mathrm{d}\theta^2 + \mathrm{d}h_1^2) + \mathrm{d}\rho_1^2,$$

and where γ itself in these coordinates is $\gamma(\theta) = (0, \theta, 0)$ and $(0, \theta, h_1)$ is at a distance of h_1 from ∂M. Choose δ so that the region $\rho_1 \in [-\delta, \delta]$ and $h_1 \geqslant -\delta$ is an embedded neighborhood of γ. We also want $\sinh(\delta) \leqslant 1$, for a reason that we will discuss later. We use these coordinates to glue M to N_1, with its metric provided by Lemma 4.2.

To glue N_1 to N_2, we introduce the coordinates (x_1, y_1, z_1) with $x_1, z_1 \in \mathbb{R}$ and $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, and with the hyperbolic metric

$$\mathrm{d}s^2 = \frac{\mathrm{d}x_1^2 + \mathrm{d}y_1^2 + \mathrm{d}z_1^2}{z_1^2}.$$

As in the proof of Lemma 4.1, we can identify these coordinates with both (ρ_1, θ, h_1) and (ρ_2, θ, h_2) using the equations

$$(x_1, y_1, z_1) = (e^{c-h_1}\tanh(\rho_1), b\theta, e^{c-h_1}\operatorname{sech}(\rho_1))$$
$$(x_1, y_1, z_1) = (\rho_2, b\theta, e^{-h_2}).$$

Both changes of variables preserve the defined metrics. We can also solve for (ρ_2, h_2) in terms of (ρ_1, h_1) to obtain

$$(\rho_2, h_2) = (e^{c-h_1}\tanh(\rho_1), h_1 - c - \log(\operatorname{sech}(\rho_1))).$$

If $\sinh(\delta) \leqslant 1$, the constant curvature end of N_1 stays within the constant curvature end of N_2, and is inserted between its corners, as in Figure 6.

Finally we define the lid L to be the region

$$L = \{(x_2, y_2, h_2) \mid x_2^2 + y_2^2 \leqslant b^2, 2 \leqslant h \leqslant 3\}$$

in $\mathbb{R}^3$ with metric

$$\mathrm{d}s^2 = e^{2h_2}(\mathrm{d}x_2^2 + \mathrm{d}y_2^2) + \mathrm{d}h_2^2.$$

We glue N_2 to L by changing to polar coordinates,

$$(x_2, y_2, h_2) = ((\rho_2 + b)\cos(\theta), (\rho_2 + b)\sin(\theta), h_2),$$

which also converts between the metrics on L and N_2.

By construction, $N_1 \cup N_2 \cup L$ is a 2-handle that attaches to M along γ. The result is a $(-1-\varepsilon, -1+\varepsilon)$-pinched 3-ball

$$\Omega = M \cup N_1 \cup N_2 \cup L$$

with piecewise smooth boundary; we can make the boundary smooth by shaving it slightly. Since M itself has arbitrarily large volume, Ω does too. It remains to bound the surface area

$$|\partial\Omega| \leqslant |\partial M| + |\partial N_1| + |\partial N_2| + |\partial L|.$$

The submanifolds N_1 and L only depends on ε. Meanwhile we have already specified that $|\partial M| \leqslant 1$. Finally $|\partial N_1|$ is estimated in Lemma 4.2 in terms of the constants a and b. The constant a is bounded because $|\gamma| \leqslant 1$, while the constant b only depends on ε. Thus $|\partial\Omega|$ is bounded by a constant, depending on ε. □

5. GEODESIC INTEGRALS

In this section we will study Santaló's integral formula [San04, Sec. 19.4] in the formalism of geodesic flow and symplectic geometry. See McDuff and Salamon [MS98, Sec. 5.4] for properties of symplectic quotients. The formulas we derive are those of Croke [Cro84]; see also Teufel [Teu93].

5.1. **Symplectic geometry.**

Let W be an open symplectic $2n$-manifold with a symplectic form ω_W. Then W also has a canonical volume form $\mu_W = \omega_W^{\wedge n}$ which is called the Liouville measure on W. Let $h : W \to \mathbb{R}$ be a Hamiltonian, by definition any smooth function on W, suppose that 0 is a regular value of h, and let $H = h^{-1}(0)$ then be the corresponding smooth level surface. Then ω_W converts the 1-form $\mathrm{d}h$ to a vector field ξ which is tangent to H. Suppose that every orbit γ of ξ only exists for a finite time interval. Let G be the set of orbits of ξ on H; it is a type of *symplectic quotient* of W. G is a smooth open manifold except that it might not be Hausdorff.

The manifold G is also symplectic with a canonical 2-form ω_G and its own Liouville measure μ_G. H cannot be symplectic since it is odd-dimensional, but it does have a Liouville measure μ_H. (In fact G and ω_G only depend on H, and not otherwise on h, while μ_H depends on the specific choice of h.) Let $(a(\gamma), b(\gamma))$ be the time interval of existence of $\gamma \in G$; here only the difference

$$\ell(\gamma) = b(\gamma) - a(\gamma)$$

is well-defined by the geometry. In this general setting, if $f : H \to \mathbb{R}$ is a suitably integrable function, then

$$\int_H f(x)\,\mathrm{d}\mu_H(x) = \int_{\gamma \in G} \int_{a(\gamma)}^{b(\gamma)} f(\gamma(t))\,\mathrm{d}t\,\mathrm{d}\mu_G(\gamma). \tag{22}$$

Or, if σ is a measure on H, we can consider the push-forward $(\pi_G)_*(\sigma)$ of σ under the projection $\pi_G : H \to G$. Taking the special case that f is constant on orbits of ξ, the relation (22) says that

$$(\pi_G)_*(\mu_H) = \ell \mu_G.$$

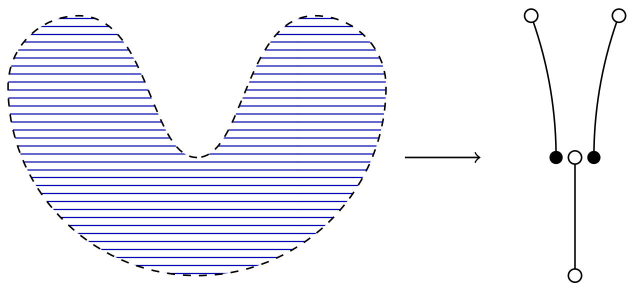

FIGURE 7. A manifold M in which the space of geodesics is not Hausdorff. The horizontal chords make a "zipper" 1-manifold.

5.2. **The space of geodesics and étendue.** If M is a smooth n-manifold, then $W = T^*M$ is canonically a symplectic manifold. If M has a Riemannian metric g, then g gives us a canonical identification $TM \cong T^*M$. It also gives us a Hamiltonian $h : TM \to \mathbb{R}$ defined as

$$h(v) = (g(v,v) - 1)/2.$$

The level surface $h^{-1}(0)$ is evidently the unit tangent bundle UM. It is less evident, but still routine, that the Hamiltonian flow ξ of h is the geodesic flow on UM. Suppose further that M only has bounded-time geodesics. Then the corresponding symplectic quotient G is the space of oriented geodesics on M. The structure on G that particularly interests us is its Liouville measure μ_G. The Liouville measure on $H = UM$ is also important, and happens to equal the Riemannian measure ν_{UM}. Even in this special case, G might not be Hausdorff if the geodesics of M merge or split, as in Figure 7.

The Liouville measure μ_G is important in geometric optics [Smi07], where among other names it is called *étendue*[1]. Lagrange established that étendue is conserved. Mathematically, this says exactly that the $(2n-2)$-form μ_G, which is definable on H, descends to G. More explicitly, suppose (in the full generality of Section 5.1) that $K_1, K_2 \subseteq H$ are two transverse open disks that are identified by the holonomy map

$$\phi : K_1 \xrightarrow{\cong} K_2$$

induced by the set of orbits. Then the Liouville measures K_1 and K_2 match, *i.e.*, $\phi_*(\mu_G) = \mu_G$. (As in the proof of Liouville's theorem, ϕ is even a symplectomorphism.)

Now suppose that Ω is a compact Riemannian manifold with boundary and with unique geodesics, and let M be the interior of Ω. Then G, the space of oriented geodesics of Ω or M, is canonically identified in two ways to $U^+\partial\Omega$. We can let $\gamma = \gamma_u$ be the geodesic generated by u, or we can let $\gamma = \overline{\gamma_u}$ be the geodesic with the same image but inverse orientation. These are both examples of identifying part of G, in this case all of G, with transverse submanifolds as in the previous paragraph. Let

$$\sigma_+ : U^+\partial\Omega \xrightarrow{\cong} G, \qquad \sigma_- : U^+\partial\Omega \xrightarrow{\cong} G$$

[1]In English, not just in French.

be the two corresponding identifications.

The maps $\sigma_\pm$ are smooth bijections; when $\partial\Omega$ is convex, they are diffeomorphisms. In general, the inverses $\sigma_\pm^{-1}$ are smooth away from the non-Hausdorff points of G. These correspond to geodesics tangent to $\partial\Omega$, and they are a set of measure 0 in G. Thus, the composition

$$\phi = \sigma_-^{-1} \circ \sigma_+ : U^+\partial\Omega \xrightarrow{\cong} U^+\partial\Omega$$

is an involution of $U^+\partial\Omega$ that preserves the measure μ_G. (It is even almost everywhere a local symplectomorphism with respect to ω_G.) We call the map ϕ the *optical transport* of Ω.

We define several types of coordinates on G, $U\Omega$ and $U^+\partial\Omega$. Let $u = (x,v)$ be the position and vector components of a tangent vector $u \in U\Omega$, and let $u = (p,v)$ be the same for $u \in U^+\partial\Omega$. On G itself, we already have the length function $\ell(\gamma)$. In addition, if $\gamma = \gamma_u$ for

$$u = (p,v) \in U^+\partial\Omega,$$

let $\alpha(\gamma)$ be the angle between v and the inward normal vector $w(p)$. If $\gamma = \overline{\gamma_u}$, then let $\beta(\gamma)$ be that angle instead.

The map σ_+ relates the Liouville measure μ_G with Riemannian measure $\nu_{U^+\partial\Omega}$. More loosely, the projection π_G from Section 5.1 relates μ_G with $\nu_{U\Omega}$. Then by slight abuse of notation,

$$d\mu_G = \cos(\alpha)\, d\nu_{U^+\partial\Omega} = \frac{d\nu_{U\Omega}}{\ell}. \tag{23}$$

In words, μ_G is close to $\nu_{U^+\partial\Omega}$ but not the same: If a beam of light is incident to a surface at an angle of α, then its illumination has a factor of $\cos(\alpha)$. The measure $(\pi_G)_*(\nu_{U\Omega})$ is also close but not the same, because the étendue of a family of geodesics does not grow with the length of the geodesics.

Another important comparison of measures relates geodesics to pairs of points.

Lemma 5.1. *Suppose that $p,q \in \Omega$ lie on a geodesic γ and that $p \neq q$. Let $p = \gamma(a)$ and $q = \gamma(b)$. Then*

$$d\nu_{\Omega\times\Omega}(p,q) = j_\Omega(\gamma,a,b)\, d\mu_G(\gamma)\, da\, db.$$

In the cases $(p,q) \in \Omega\times\partial\Omega$, $(p,q) \in \partial\Omega\times\Omega$ and $(p,q) \in \partial\Omega\times\partial\Omega$ we respectively have

$$d\nu_{\Omega\times\partial\Omega}(p,q) = \frac{j_\Omega(\gamma,a,b)}{\cos\beta(\gamma)}\, d\mu_G(\gamma)\, da$$

$$d\nu_{\partial\Omega\times\Omega}(p,q) = \frac{j_\Omega(\gamma,a,b)}{\cos\alpha(\gamma)}\, d\mu_G(\gamma)\, db$$

$$d\nu_{\partial\Omega\times\partial\Omega}(p,q) = \frac{j_\Omega(\gamma,a,b)}{\cos\alpha(\gamma)\cos\beta(\gamma)}\, d\mu_G(\gamma).$$

Proof. We start with the case $p,q \in \Omega$. On the one hand, a localized version of formula (23) is

$$d\mu_G(\gamma)\, da = d\nu_{U\Omega}(u) = d\nu_{U_p\Omega}(v)\, d\nu_\Omega(p)$$

when $u = (p,v)$ and γ is the geodesic such that $\gamma(a) = p$ and $\gamma'(a) = v$. On the other hand, by the definition of the candle function, we have for all fixed p:

$$d\nu_\Omega(q) = j_\Omega(\gamma,a,b)\, d\nu_{U_p\Omega}(v)\, db$$

when $\gamma(a) = p$, $\gamma(b) = q$ and $v = \gamma'(a)$. Thus, we obtain a pair of equalities of measures:

$$d\mu_G(\gamma)\, da\, db = d\nu_\Omega(p)\, d\nu_{U_p\Omega}(v)\, db = \frac{d\nu_\Omega(p)\, d\nu_\Omega(q)}{j_\Omega(\gamma,a,b)}.$$

The other cases are handled in the same way. When $p \in \Omega$ and $q \in \partial\Omega$ we have

$$\begin{aligned} \mathrm{d}\mu_G(\gamma)\,\mathrm{d}a &= \mathrm{d}\nu_{U_p\Omega}(v)\,\mathrm{d}\nu_\Omega(p) \\ \mathrm{d}\nu_{\partial\Omega}(q) &= \frac{j_\Omega(\gamma,a,b)}{\cos\beta(\gamma)}\,\mathrm{d}\nu_{U_p\Omega}(v) \end{aligned}$$

where b is entirely determined by p and v, instead of being a variable; when $p \in \partial\Omega$ and $q \in \Omega$ then a can be fixed to 0 by choosing a suitable parametrization of geodesics, and we have

$$\begin{aligned} \mathrm{d}\mu_G(\gamma)\,\mathrm{d}b &= \cos\alpha(\gamma)\,\mathrm{d}\nu_{U_p^+\partial\Omega}(v)\,\mathrm{d}\nu_{\partial\Omega}(p)\,\mathrm{d}b \\ \mathrm{d}\nu_\Omega(q) &= j_\Omega(\gamma,a,b)\,\mathrm{d}\nu_{U_p^+\partial\Omega}(v)\,\mathrm{d}b \end{aligned}$$

and finally when $(p,q) \in \partial\Omega \times \partial\Omega$ we use

$$\begin{aligned} \mathrm{d}\mu_G(\gamma) &= \cos\alpha(\gamma)\,\mathrm{d}\nu_{U_p^+\partial\Omega}(v)\,\mathrm{d}\nu_{\partial\Omega}(p) \\ \mathrm{d}\nu_{\partial\Omega}(q) &= \frac{j_\Omega(\gamma,a,b)}{\cos\beta(\gamma)}\,\mathrm{d}\nu_{U_p^+\partial\Omega}(v). \end{aligned} \qquad \square$$

Lemma 5.1 has the important corollary that the candle function is symmetric. To generalize from an Ω with unique geodesics to an arbitrary M, we can let Ω be a neighborhood of the geodesic γ, immersed in M.

Corollary 5.2 (Folklore [Yau75, Lem. 5])**.** *In any Riemannian manifold M,*

$$j_M(\gamma,a,b) = j_M(\gamma,b,a).$$

Combining (22) with (23) yields Santaló's equality.

Theorem 5.3 (Santaló [San04, Sec. 19.4])**.** *If Ω is as above, and if $f : U\Omega \to \mathbb{R}$ is a continuous function, then*

$$\int_{U\Omega} f(u)\mathrm{d}\nu_{U\Omega}(u) = \int_{U^+\partial\Omega}\int_0^{\ell(\gamma_u)} f(\gamma_u(t))\cos(\alpha(u))\,\mathrm{d}t\,\mathrm{d}\nu_{U^+\partial\Omega}(u).$$

Finally, we will consider another reduction of the space G, the projection

$$\pi_{\mathrm{lab}} : G \to \mathbb{R}_{\geqslant 0} \times [0,\pi/2]^2, \quad \pi_{\mathrm{lab}}(\gamma) = (\ell(\gamma),\alpha(\gamma),\beta(\gamma)).$$

Let

$$\mu_\Omega = (\pi_{\mathrm{lab}})_*(\mu_G)$$

be the push-forward of Liouville measure. Then μ_Ω is a measure-theoretic reduction of the optical transport map ϕ, and is close to a transportation measure in the sense of Monge-Kantorovich. More precisely, equation (23) yields a formula for the α and β marginals of μ_Ω, so we can view μ_Ω, or rather its projection to $[0,\pi/2)^2$, as a transportation measure from one marginal to the other. The projection onto the α coordinate is

$$\alpha_*(\mu_\Omega) \overset{\mathrm{def}}{=} \int_{\ell,\beta} \mathrm{d}\mu_\Omega = |\partial\Omega|\omega_{n-2}\sin(\alpha)^{n-2}\cos(\alpha)\,\mathrm{d}\alpha,$$

where as in (12) we use the volume of a latitude sphere on $U_p^+\partial\Omega$. Using the abbreviation

$$z(\theta) = \frac{\omega_{n-2}\sin(\theta)^{n-1}}{n-1},$$

we can give a simplified formula for both marginals:

$$\alpha_*(\mu_\Omega) = |\partial\Omega|\,\mathrm{d}z(\alpha), \qquad \beta_*(\mu_\Omega) = |\partial\Omega|\,\mathrm{d}z(\beta). \tag{24}$$

A final important property of μ_Ω that follows from its construction is that it is symmetric in α and β.

5.3. **The core inequalities.** In this section, we establish three geometric comparisons that convert our curvature hypotheses to linear inequalities that can then be used for linear programming. (Section 5.2 does not use either unique geodesics or a curvature hypothesis. Thus, the results there are not strong enough to establish an isoperimetric inequality.)

Lemma 5.4. *If Ω is* Candle(κ) *and has unique geodesics, then:*

$$\int_{\ell,\alpha,\beta} \frac{s_{n,\kappa}(\ell)}{\cos(\alpha)\cos(\beta)}\,\mathrm{d}\mu_\Omega \leqslant |\partial\Omega|^2 \qquad \textit{(Croke)} \tag{25}$$

$$\int_{\ell,\alpha,\beta} \frac{s^{(-1)}_{n,\kappa}(\ell)}{\cos(\alpha)}\,\mathrm{d}\mu_\Omega \leqslant |\partial\Omega||\Omega| \qquad \textit{(Little Prince)} \tag{26}$$

$$\int_{\ell,\alpha,\beta} s^{(-2)}_{n,\kappa}(\ell)\,\mathrm{d}\mu_\Omega \leqslant |\Omega|^2 \qquad \textit{(Teufel)} \tag{27}$$

If Ω is convex and has constant curvature κ, then all three inequalities are equalities.

The first case of Lemma 5.4, equation (25), is due to Croke [Cro84]. Equation (26) generalizes the integral over $p \in \partial\Omega$ of equation (9) in Theorem 1.6. Finally equation (27) generalizes an isoperimetric inequality of Teufel [Teu91]. Nonetheless all three inequalities can be proven in a similar way.

Proof. We define a *partial* map $\tau : \Omega \times \Omega \to G$ by letting $\tau(p,q)$ be the unique geodesic $\gamma \in G$ that passes through p and q, if it exists. We define $\tau(p,q)$ only when $p \neq q$ and only when γ is available. Also, if γ exists, we parametrize it by length starting at the initial endpoint at 0.

By construction,

$$\begin{aligned} ||\tau_*(\nu_{\partial\Omega\times\partial\Omega})|| &\leqslant |\partial\Omega|^2 \\ ||\tau_*(\nu_{\partial\Omega\times\Omega})|| &\leqslant |\partial\Omega||\Omega| \\ ||\tau_*(\nu_{\Omega\times\Omega})|| &\leqslant |\Omega|^2. \end{aligned}$$

Note that each inequality is an equality if and only if Ω is convex.

Using Lemma 5.1, we can write integrals for each of the left sides

$$\begin{aligned} ||\tau_*(\nu_{\partial\Omega\times\partial\Omega})|| &= \int_G \frac{j_\Omega(\gamma,0,\ell)}{\cos(\alpha)\cos(\beta)}\,\mathrm{d}\mu_G(\gamma) \\ ||\tau_*(\nu_{\partial\Omega\times\Omega})|| &= \int_G \int_0^\ell \frac{j_\Omega(\gamma,0,r)}{\cos(\alpha)}\,\mathrm{d}r\,\mathrm{d}\mu_G(\gamma) \\ ||\tau_*(\nu_{\Omega\times\Omega})|| &= \int_G \int_0^\ell \int_0^t j_\Omega(\gamma,r,t)\,\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\mu_G(\gamma). \end{aligned}$$

Because Ω is Candle(κ),

$$\begin{aligned} j_\Omega(\gamma,0,\ell) &\geqslant s_{n,\kappa}(\ell), \\ \int_0^\ell j_\Omega(\gamma,0,t)\,\mathrm{d}t &\geqslant s^{(-1)}_{n,\kappa}(\ell), \\ \int_0^\ell \int_0^t j_\Omega(\gamma,r,t)\,\mathrm{d}r\,\mathrm{d}t &\geqslant s^{(-2)}_{n,\kappa}(\ell), \end{aligned}$$

and note that each inequality is an equality if Ω has constant curvature κ. We thus obtain

$$\int_G \frac{s_{n,\kappa}(\ell)}{\cos(\alpha)\cos(\beta)}\, d\mu_G(\gamma) \leqslant |\partial\Omega|^2$$
$$\int_G \frac{s_{n,\kappa}^{(-1)}(\ell)}{\cos(\alpha)}\, d\mu_G(\gamma) \leqslant |\partial\Omega||\Omega|$$
$$\int_G s_{n,\kappa}^{(-2)}(\ell)\, d\mu_G(\gamma) \leqslant |\Omega|^2.$$

Because these integrands only depend on ℓ, α, and β, we can now descend from μ_G to μ_Ω. □

5.4. **Extended inequalities.** Lemma 5.4 will yield a linear programming model that is strong enough to prove Theorem 1.4, but not Theorem 1.5 nor many of the other cases of Theorem 1.17. In this section, we will establish several variations of Lemma 5.4 using alternate hypotheses.

The following lemma is the refinement needed for Theorem 1.5 and Theorem 1.16.

Lemma 5.5. *Suppose that Ω is a compact domain in an* LCD(-1) *Cartan-Hadamard n-manifold M and let*

$$\mathrm{chord}(\Omega) \leqslant L \in (0,\infty].$$

Then

$$\int_{\ell,\alpha,\beta} \Big(\frac{s_{n,-1}^{(-1)}(\ell)}{\cos(\alpha)} - \frac{(n-1)s_{n,-1}^{(-2)}(\ell)}{\tanh(L)}\Big)\, d\mu_\Omega \leqslant |\partial\Omega||\Omega| - \frac{(n-1)|\Omega|^2}{\tanh(L)} \tag{28}$$

$$\int_{\ell,\alpha,\beta} \Big(\frac{s_{n,-1}(\ell)}{\cos(\alpha)\cos(\beta)} - \frac{(n-1)s_{n,-1}^{(-1)}(\ell)}{\tanh(L)\cos(\alpha)}\Big)\, d\mu_\Omega \leqslant |\partial\Omega|^2 - \frac{(n-1)|\partial\Omega||\Omega|}{\tanh(L)}. \tag{29}$$

If ω is convex and has constant curvature κ, then the inequalities are equalities.

Proof of (28). We abbreviate

$$s(\ell) \stackrel{\text{def}}{=} s_{n,-1}(\ell),$$

and we switch α and β in the integral.

Let G be the space of geodesics of Ω and recall the partial map $\tau : \Omega\times\Omega \to G$ used in the proof of Lemma 5.4 and the measures $\nu_{\partial\Omega\times\Omega}$ and $\nu_{\Omega\times\Omega}$. We consider the signed measure

$$\sigma_{\Omega\times\Omega} \stackrel{\text{def}}{=} \nu_{\Omega\times\partial\Omega} - \frac{n-1}{\tanh(L)}\nu_{\Omega\times\Omega}.$$

To be precise, if $(p,q) \in \Omega\times\partial\Omega$, then $\gamma = \tau(p,q)$ is the geodesic that passes through p and ends at q. We claim two things about the pushforward $\tau_*(\sigma_{\Omega\times\Omega})$:

1. That the net measure omitted by τ is non-negative:

$$||\tau_*(\sigma_{\Omega\times\Omega})|| \leqslant |\partial\Omega||\Omega| - \frac{n-1}{\tanh(L)}|\Omega|^2.$$

2. That the measure that is pushed forward is underestimated by the comparison candle function:

$$\int_G \Big(\frac{s^{(-1)}(\ell)}{\cos(\beta)} - \frac{(n-1)s^{(-2)}(\ell)}{\tanh(L)}\Big)\, d\mu_G(\gamma) \leqslant ||\tau_*(\sigma_{\Omega\times\Omega})||.$$

Just as in the proof of Lemma 5.4, equation (28) follows from these two claims.

To prove the second claim, let $\gamma \in G$ be a maximal geodesic of Ω with unit speed and domain $[0,\ell]$. We abbreviate the candle function along γ:

$$j(t) \stackrel{\text{def}}{=} j(\gamma,t), \qquad j(r,t) \stackrel{\text{def}}{=} j(\gamma,r,t).$$

Since M and therefore Ω is $\mathrm{LCD}(-1)$, we have the inequality

$$\frac{j'(t)}{j(t)} \geqslant \frac{s'(t)}{s(t)}.$$

We can rephrase this as saying that

$$\frac{\partial j}{\partial t}(r,t) - \frac{s'(t-r)}{s(t-r)} j(r,t) = \frac{\partial j}{\partial t}(r,t) - \frac{n-1}{\tanh(t-r)} j(r,t)$$

is minimized (with a value of 0) in the $K=-1$ case. Now

$$\tanh(t-r) \leqslant \tanh(L),$$

while $\mathrm{LCD}(-1)$ implies $\mathrm{Candle}(-1)$, *i.e.*,

$$j(r,t) \geqslant s(t-r).$$

It follows that

$$\frac{\partial j}{\partial t}(r,t) - \frac{(n-1)j(r,t)}{\tanh(L)} \geqslant s'(t-r) - \frac{(n-1)s(t-r)}{\tanh(L)}. \tag{30}$$

We can integrate with respect to r and t to obtain:

$$\begin{aligned}\int_0^\ell \int_r^\ell \Big[\frac{\partial j}{\partial t}(r,t) - \frac{(n-1)j(r,t)}{\tanh(L)}\Big]\,\mathrm{d}t\,\mathrm{d}r &= \int_0^\ell j(r,\ell)\,\mathrm{d}r - \frac{n-1}{\tanh(L)}\int_0^\ell \int_r^\ell j(r,t)\,\mathrm{d}r\,\mathrm{d}t \\ &\geqslant s^{(-1)}(\ell) - \frac{(n-1)s^{(-2)}(\ell)}{\tanh(L)}.\end{aligned}$$

Then, if the terminating angle of γ is β, we can again use the $\mathrm{Candle}(-1)$ condition to obtain

$$\int_0^\ell \frac{j(r,\ell)}{\cos(\beta)}\,\mathrm{d}r - \frac{n-1}{\tanh(L)}\int_0^\ell \int_0^t j(r,t)\,\mathrm{d}r\,\mathrm{d}t \geqslant \frac{s^{(-1)}(\ell)}{\cos(\beta)} - \frac{(n-1)s^{(-2)}(\ell)}{\tanh(L)}.$$

Since the left side is the fiber integral of $\tau_*(\sigma_{\Omega\times\Omega})$, as in the proof of Lemma 5.4, this establishes the second claim.

To establish the first claim, for each $p \in \Omega$, we consider the set $\Omega \setminus \mathrm{Vis}(\Omega,p)$ consisting of points $q \in \Omega$ that are *not* visible from p. The union of all of these is exactly the pairs (p,q) where τ is not defined. If γ is a geodesic in M emanating from p, we can restrict further to its intersection

$$\gamma \cap (\Omega \setminus \mathrm{Vis}(\Omega,p)),$$

where we extend the geodesic γ from Ω to M. We claim that the integral of $\sigma_{\Omega\times\Omega}$ on each of these intersections, with the appropriate Jacobian factor, is non-negative.

To verify this claim, we suppose that the intersection is non-empty, and we parametrize γ at unit speed so that $\gamma(0) = p$. Let I be the set of times t such that

$$\gamma(t) \in \Omega \setminus \mathrm{Vis}(\Omega,p),$$

let $\{t_k\}$ be the set of right endpoints of I where γ leaves Ω, and for each k, let $\beta_k \in [0,\pi/2]$ be the angle that γ exits Ω at $\gamma(t_k)$. Let ℓ be the rightmost point of I. Then the infinitesimal portion of $\sigma_{\Omega\times\Omega}$ on $\gamma(I)$ is

$$\sum_k \frac{j(t_k)}{\cos(\beta_k)} - \frac{n-1}{\tanh(L)} \int_I j(t)\,\mathrm{d}t \geqslant j(\ell) - \frac{n-1}{\tanh(L)} \int_0^\ell j(t)\,\mathrm{d}t.$$

(In other words, we geometrically simplify to the worst case: $I = [0,\ell]$ and $\beta = 0$.) The derivative of the right side is now

$$j'(\ell) - \frac{n-1}{\tanh(L)} j(\ell) \geqslant 0. \tag{31}$$

The inequality holds because it is the same as (30), except with the right side simplified to 0. This establishes the first claim and thus (28).

The equality criterion holds for the same reasons as in Lemma 5.4. □

Proof of (29). The proof has exactly the same ideas as the proof of (28), only with some changes to the formulas. We keep the same abbreviations. This time we define

$$\sigma_{\partial\Omega\times\Omega} \stackrel{\text{def}}{=} \nu_{\partial\Omega\times\partial\Omega} - \frac{n-1}{\tanh(L)} \nu_{\partial\Omega\times\Omega},$$

we consider $\tau_*(\sigma_{\partial\Omega\times\Omega})$, and we claim:

1. That the net measure omitted by τ is non-negative:

$$||\tau_*(\sigma_{\partial\Omega\times\Omega})|| \leqslant |\partial\Omega|^2 - \frac{n-1}{\tanh(L)} |\partial\Omega||\Omega|^2.$$

2. That the integral underestimates the pushforward:

$$\int_G \Big(\frac{s(\ell)}{\cos(\alpha)\cos(\beta)} - \frac{(n-1)s^{(-1)}(\ell)}{\cos(\alpha)\tanh(L)} \Big)\,\mathrm{d}\mu_G(\gamma) \leqslant ||\tau_*(\sigma_{\partial\Omega\times\Omega})||.$$

To prove the second claim, we define γ and j as before and we again obtain (30). In this case, we integrate only with respect to $t \in [0,\ell]$ to obtain

$$j(0,\ell) - \frac{n-1}{\tanh(L)} \int_0^\ell j(0,t)\,\mathrm{d}t \geqslant s(\ell) - \frac{(n-1)s^{(-1)}(\ell)}{\tanh(L)}.$$

Now divide through by $\cos(\alpha)$, and we use the Candle(-1) property to divide the first term $\cos(\beta)$, to obtain

$$\frac{j(0,\ell)}{\cos(\alpha)\cos(\beta)} - \frac{n-1}{\cos(\alpha)\tanh(L)} \int_0^\ell j(0,t)\,\mathrm{d}t \geqslant \frac{s(\ell)}{\cos(\alpha)\cos(\beta)} - \frac{(n-1)s^{(-1)}(\ell)}{\cos(\alpha)\tanh(L)}.$$

The left side is the fiber integral of $\tau_*(\sigma_{\partial\Omega\times\Omega})$, so this establishes the second claim.

The proof of the first claim is identical to the case of (28), except that $p \in \partial\Omega$, and we divide through by $\cos(\alpha)$. □

Meanwhile Theorem 1.15 requires the following striking inequality that depends only on the condition of unique geodesics rather than any bound on curvature. We omit the proof as the lemma is equivalent to Lemma 9 of Croke [Cro80].

Lemma 5.6 (Croke-Berger-Kazdan)**.** *If Ω is a compact Riemannian manifold with boundary and with unique geodesics, then*

$$\int_{\ell,\alpha,\beta} s^{(-2)}_{n,(\pi/\ell)^2}(\ell)\,\mathrm{d}\mu_\Omega \leqslant |\Omega|^2.$$

5.5. **Mirrors and multiple images.** In this section, we establish the geometric inequalities needed for Theorem 1.8. Let M be a Riemannian manifold with boundary ∂M (although M might not be compact), and consider geodesics that reflect from ∂M with equal angle of incidence and angle of reflection. We then have an extension of the exponential map at any point $x \in M$ beyond the time of reflection. This extended exponential map has a natural Jacobian; thus M has an extended candle function $j_M(\gamma, r)$ as treated in Section 2.2. Then we say that M is $\mathrm{Candle}(\kappa)$ in the sense of reflecting geodesics if this Jacobian satisfies the $\mathrm{Candle}(\kappa)$ comparison.

Let $\hat{G}$ be the space of these geodesics, for simplicity considering only those geodesics that are never tangent to ∂M. Then the results of Section 5.2 still apply, with only slight modifications. In particular M might have a compactification Ω with $\partial M = W \subseteq \partial\Omega$. Then (23) applies if we replace $\partial\Omega$ by $\partial\Omega \setminus W$; Lemma 5.1 holds; etc.

If Ω has a mirror W as part of its boundary, then some pairs of points have at least two connecting, reflecting geodesics. We can suppose in general that every two points in (Ω, W) are connected by at most m geodesics (which is also interesting even if W is empty), and we can suppose that (Ω, W) is $\mathrm{Candle}(\kappa)$ in the sense of reflecting geodesics. In this case it is straightforward to generalize Lemma 5.4. The generalization will yield the linear programming model for Theorem 1.8.

Lemma 5.7. *If* (Ω, W) *is* $\mathrm{Candle}(\kappa)$ *and has at most* m *reflecting geodesics between any pair of points, then:*

$$\int_{\alpha,\beta,\ell} \frac{s_{n,\kappa}(\ell)}{\cos(\alpha)\cos(\beta)}\,\mathrm{d}\mu_\Omega \leqslant m|\partial\Omega|^2 \tag{32}$$

$$\int_{\alpha,\beta,\ell} \frac{s_{n,\kappa}^{(-1)}(\ell)}{\cos(\alpha)}\,\mathrm{d}\mu_\Omega \leqslant m|\partial\Omega||\Omega| \tag{33}$$

$$\int_{\alpha,\beta,\ell} s_{n,\kappa}^{(-2)}(\ell)\,\mathrm{d}\mu_\Omega \leqslant m|\Omega|^2. \tag{34}$$

Proof. The proof is nearly identical to that of Lemma 5.4. In this case

$$\tau : \Omega \times \Omega \to G$$

is not a partial map, but rather a multivalued correspondence which is at most 1 to m everywhere. We can define a pushforward measure such as $\tau_*(\nu_{\Omega\times\Omega})$ by counting multiplicities.

By construction:

$$||\tau_*(\nu_{\partial\Omega\times\partial\Omega})|| \leqslant m|\partial\Omega|^2$$
$$||\tau_*(\nu_{\partial\Omega\times\Omega})|| \leqslant m|\partial\Omega||\Omega|$$
$$||\tau_*(\nu_{\Omega\times\Omega})|| \leqslant m|\Omega|^2.$$

On the other hand,

$$||\tau_*(\nu_{\partial\Omega\times\partial\Omega})|| = \int_G \frac{j(\gamma,\ell)}{\cos(\alpha)\cos(\beta)}\,\mathrm{d}\mu_G(\gamma)$$
$$||\tau_*(\nu_{\partial\Omega\times\Omega})|| = \int_G \int_0^\ell \frac{j(\gamma,r)}{\cos(\alpha)}\,\mathrm{d}r\,\mathrm{d}\mu_G(\gamma)$$
$$||\tau_*(\nu_{\Omega\times\Omega})|| = \int_G \int_0^\ell \int_0^t j(\gamma,r,t)\,\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}\mu_G(\gamma).$$

Using the $\mathrm{Candle}(\kappa)$ hypothesis, we obtain the desired inequalities. $\square$

Finally, the following generalization of Günther's inequality [Gün60, BC64] shows that the Candle(κ) condition is actually useful for reflecting geodesics.

Proposition 5.8. *Let M be a Riemannian manifold with $K \leqslant \kappa$ for some $\kappa \in \mathbb{R}$, and suppose that ∂M is concave relative to the interior. If $\kappa > 0$, suppose also that* $\mathrm{chord}(M) < \pi/\sqrt{\kappa}$*. Then M is* $\mathrm{LCD}(\kappa)$ *with respect to geodesics that reflect from ∂M.*

Proposition 5.8 generalizes Lemma 3.2 of Choe [Cho06], which claims Candle(0) using (in the proof) the same hypotheses when $\kappa = 0$. However, the argument given there omits many details about reflection from a convex surface. (Which is thus concave from the other side as we describe it.) We will prove Proposition 5.8 in Section 8.2.

6. Linear programming and optimal transport

6.1. A linear model for isoperimetric problems.

6.1.1. *Feasibility.* In this section we abstract the results of Section 5.2 and 5.3 into a linear programming model.

Assume that Ω is an n-manifold with boundary, with unique geodesics, and with curvature at most κ. Let $V = |\Omega|$ and $A = |\partial\Omega|$. Then equations (23), (24) (25), (26), and (27) show (after symmetrization in α and β) that $\mu = \mu_\Omega$ is a solution to the following infinite linear programming problem.

LP Problem 6.1. *Given n, κ, A, and V, let*

$$z(\theta) = \frac{\omega_{n-2}\sin(\theta)^{n-1}}{n-1}.$$

Is there a positive measure $\mu(\ell,\alpha,\beta)$ on $\mathbb{R}_{\geqslant 0}\times[0,\pi/2)^2$, which is symmetric in α and β, and such that

$$\alpha_*(\mu) = \int_{\ell,\beta} \mathrm{d}\mu = A\,\mathrm{d}z(\alpha) \tag{35}$$

$$\int_{\ell,\alpha,\beta} s_{n,\kappa}(\ell)\sec(\alpha)\sec(\beta)\,\mathrm{d}\mu \leqslant A^2 \tag{36}$$

$$\int_{\ell,\alpha,\beta} s^{(-1)}_{n,\kappa}(\ell)\big(\sec(\alpha)+\sec(\beta)\big)\,\mathrm{d}\mu \leqslant 2AV \tag{37}$$

$$\int_{\ell,\alpha,\beta} s^{(-2)}_{n,\kappa}(\ell)\,\mathrm{d}\mu \leqslant V^2 \tag{38}$$

$$\int_{\ell,\alpha,\beta} \ell\,\mathrm{d}\mu = \omega_{n-1}V? \tag{39}$$

(We could have written Problem 6.1 without symmetrization in α and β. It would have been equivalent, but more complicated.)

Since our ultimate goal is to prove a lower bound for $|\partial\Omega|$, we want to show that given n, κ, and V, Problem 6.1 is infeasible for values of A that are too low. As usual in linear programming, we will profit from stating a dual problem.

LP Problem 6.2 (Dual to Problem 6.1)**.** *Given n, κ, A, and V, are there numbers $a,b,c \geqslant 0$ and $d \in \mathbb{R}$ and a continuous function $f : [0,\pi/2) \to \mathbb{R}$ such that*

$$\begin{aligned} as_{n,\kappa}(\ell)\sec(\alpha)\sec(\beta) + bs^{(-1)}_{n,\kappa}(\ell)\big(\sec(\alpha)+\sec(\beta)\big) \\ + cs^{(-2)}_{n,\kappa}(\ell) - d\ell + f(\alpha) + f(\beta) \geqslant 0 \end{aligned} \tag{40}$$

$$aA^2 + 2bAV + cV^2 - d\,\omega_{n-1}V + 2A\int_0^{\pi/2} f(\alpha)\,\mathrm{d}z(\alpha) < 0 \tag{41}$$

for all $(\alpha,\beta,\ell) \in [0,\pi/2)^2 \times \mathbb{R}_{\geqslant 0}$*?*

(Note that the constant d can have either sign. We subtract it so that it will be positive in actual usage.)

We will discuss in what sense Problem 6.2 is dual to Problem 6.1, and the consequences of this duality, in Section 6.2. For now, we will concentrate on sufficient criteria to prove our main theorems. Problem 6.2 is strong enough to prove Theorem 1.4. In Sections 7.3 and 8, we will state other linear programming problems to handle our other results stated in Section 1.

In the rest of this section (Section 6.1), including in the statements of the lemmas, we fix V, n, and κ, but not A.

Lemma 6.3. *Let* $a,b,c \geqslant 0$*, let* $d \in \mathbb{R}$*, and let*

$$E(\ell,\alpha,\beta) = as_{n,\kappa}(\ell)\sec(\alpha)\sec(\beta) + bs^{(-1)}_{n,\kappa}(\ell)\big(\sec(\alpha)+\sec(\beta)\big) + cs^{(-2)}_{n,\kappa}(\ell) - d\ell. \quad (42)$$

Let $f : [0,\pi/2) \to \mathbb{R}$ *be a continuous function such that* $\int_0^{\pi/2} f(\alpha)\,\mathrm{d}z(\alpha)$ *is absolutely convergent. If*

$$F(\ell,\alpha,\beta) \stackrel{\text{def}}{=} E(\ell,\alpha,\beta) + f(\alpha) + f(\beta) \geqslant 0, \quad (43)$$

then Problem 6.1 is infeasible for those $A \geqslant 0$ *such that*

$$P(A) \stackrel{\text{def}}{=} aA^2 + 2bAV + cV^2 - d\omega_{n-1}V + 2A\int_0^{\pi/2} f(\alpha)\,\mathrm{d}z(\alpha) < 0.$$

Explicitly, if $P(A)$ *has two real roots, let the roots be* $A_0 < A_1$*; if* $P(A)$ *is linear, let* A_1 *be its root and let* $A_0 = -\infty$*. Then* $A \in (A_0,A_1) \cap [0,\infty)$ *is infeasible.*

We introduce some terminology which will be justified in Section 6.2.2. $E(\ell,\alpha,\beta)$ is a *cost function*, $f(\alpha)$ is a *potential*, and $F(\ell,\alpha,\beta)$ is an *adjusted cost function*.

Proof. Except for a change of variables, the proposition is the assertion that if Problem 6.2 is feasible, then Problem 6.1 is infeasible. More explicitly: For any $a,b,c \geqslant 0$, $d \in \mathbb{R}$, and suitable $f : [0,\pi/2) \to \mathbb{R}$, we can combine the relations in Problem 6.1 to produce a formula of the form

$$\int_{\ell,\alpha,\beta} F(\ell,\alpha,\beta)d\mu \leqslant P(A). \quad (44)$$

If the integrand $F(\ell,\alpha,\beta)$ is non-negative while the upper bound $P(A)$ is strictly negative, then the measure μ cannot exist. □

At this point, there is a potential difference between solving a geometric isoperimetric problem and solving a linear programming model for one. (Recall that Theorem 1.17 promises the latter.) Geometrically, the set of possible values of A in each of our isoperimetric problems must be an open or closed ray in $\mathbb{R}_+$. Thus, if we apply Lemma 6.3 to exclude $A \in (A_0,A_1)$, then all values of $A \in [0,A_1)$ are geometrically impossible even if $A_0 \geqslant 0$. Problem 6.1 has the same property because we can think of $A = A(\mu)$ as a function of μ, and we can increase A by adding measure to μ at $\ell = 0$. But this is less clear for Problem 7.2 which we will need later; in any case, a single use of Lemma 6.3 need not satisfy $A_0 < 0$. The simplest way to establish Theorem 1.17 is to obtain a negative value of A_0 and a sharp value of A_1 in Lemma 6.3. This is the case if and only if

$$P(0) = cV^2 - d\omega_{n-1}V < 0,$$

which simplifies to

$$cV < d\omega_{n-1}. \quad (45)$$

We will attain the condition (45) for Problem 6.1. In our treatment of Problem 7.2, we will take a slightly more complicated approach.

6.1.2. *Optimality.* Suppose that for some value A_1, we find a solution μ to Problem 6.1, and we find (a,b,c,d,f) in Lemma 6.3 with $P(A_1)=0$ and $P'(A_1)>0$; and suppose that (45) also holds. Then the two solutions are an *optimal pair*. The existence of (a,b,c,d,f) shows that $A=A_1$ is the smallest feasible value in Problem 6.1; the existence of μ shows that $A=A_1$ is the smallest infeasible value in Problem 6.2.

There is a simple test of whether μ and (a,b,c,d,f) are an optimal pair. If they are, then the conditions (44), $F\geqslant 0$, and $P(A)=0$ tell us that μ is supported on the zero locus of the adjusted cost F. On the other hand, if (a,b,c,d,f) satisfies both Lemma 6.3 and (45), if μ is a solution to Problem 6.1 that is supported on the zero locus of F, and if we happen to know that all inequalities in Problem 6.1 are equalities, then we can calculate that $P(A)=0$.

If in addition $\mu=\mu_\Omega$ for an admissible domain Ω, then $A=|\partial\Omega|$ is the sharp isoperimetric value. Recall that we plan to prove that $\Omega=B_{n,\kappa}$ is an isoperimetric minimizer. First, since this Ω is convex and has constant curvature κ, Lemma 5.4 tells us that all inequalities in Problem 6.1 are indeed equalities. Second, the geodesics of this Ω have the property that $\alpha=\beta$ and that $\ell=h(\alpha)$ is a function of α. If we combine these properties with the assumption that μ_Ω is part of an optimal pair and is thus supported on the zero locus of F, then we can solve for f, once we know (a,b,c,d). Together with the rest of the discussion in this section, we obtain the following sufficient criterion.

Lemma 6.4. *Suppose that $a,b,c\geqslant 0$ and $d\in\mathbb{R}$ are numbers that satisfy (45), and that $h:[0,\pi/2)\to\mathbb{R}_{\geqslant 0}$ is a continuous function. Let μ be the unique measure that satisfies (35) for some $A=A(\mu)$ and that is supported on the set $(h(\alpha),\alpha,\alpha)$, and let*

$$f(\alpha)=-\frac{E(h(\alpha),\alpha,\alpha)}{2}. \tag{46}$$

If μ is a solution to Problem 6.1, and if f satisfies (43) for the same A, then μ and (a,b,c,d,f) are an optimal pair. If in addition $\mu=\mu_\Omega$ for an admissible domain Ω, then $A=|\partial\Omega|$ is the sharp isoperimetric value.

Lemma 6.4 is the basis for our proof of Theorem 1.4 and the corresponding part of Theorem 1.17. We will use similar reasoning to prove Theorems 1.5 and 1.8. The calculations will be organized as follows. We temporarily assume the conclusion, that μ_Ω is optimal when $\Omega=B_{n,\kappa}(r)$. This yields the dependence $\ell=h(\alpha)$. If our construction were to work, the adjusted cost $F(\ell,\alpha,\beta)$ would attain a minimum of 0 at $(h(\alpha),\alpha,\alpha)$. Thus, we can solve for a, b, c, and d by applying a derivative test to the cost E or the adjusted cost F, namely,

$$\frac{\partial F}{\partial \ell}(\ell,\alpha,\alpha)=\frac{\partial E}{\partial \ell}(\ell,\alpha,\alpha)=0 \tag{47}$$

when $\ell=h(\alpha)$ and $0\leqslant\ell\leqslant 2r$.

Having calculated a, b, c, and d, which determine $E(\ell,\alpha,\beta)$, (46) tells us $f(\alpha)$. The remaining hard part of the proof is then to confirm (43). We will carry out these calculations in Section 7.

Our approach to solving Problem 6.1 as outlined in this section may seem both lucky and creative. It is indeed lucky, in the sense that (47) is an equality of functions used to solve for four numbers; it only has solutions when $n\in\{2,4\}$. In Section 6.2, we will argue that solving Problem 6.1 follows the precepts of linear programming and optimal transport with fairly little creativity.

6.2. **Generalities.** As explained at the end of Section 6.1.2, this section only provides context and is not needed for the proofs of our results.

6.2.1. *Linear programming.* The genesis of linear programming is a structure theorem for finite systems of linear equalities and inequalities due to Farkas and Minkowski [Far01, Min10, Kje02].

Theorem 6.5 (Farkas-Minkowski)**.** *Let $x = \{x_i\}$ be a finite list of real variables, and a finite system L of linear inequalities and equalities, given by two matrices A and B:*

$$\sum_i A_{j,i} x_i \leqslant a_j \qquad \sum_i B_{k,i} x_i = b_k.$$

Then:

1. *The system L is infeasible if and only if some linear combination of the form*
$$\sum_{i,j} y_j A_{j,i} x_i + \sum_{i,k} z_k B_{k,i} x_i \leqslant \sum_j y_j a_j + \sum_k z_k b_k \qquad \forall j, y_j \geqslant 0 \tag{48}$$
simplifies to the falsehood $0 \leqslant -1$. (Or $0 \leqslant c$ for some constant $c < 0$.)
2. *A linear bound $\sum_i c_i x_i \leqslant c$ holds for solutions to L if and only if it is expressible in the form* (48)*.*
3. *If L is feasible and $\sum_i c_i x_i$ is bounded on its solution set, then it has a maximum c, which is also the minimum of the right side of* (48) *subject to the constraint that the left side simplifies to $\sum_i c_i x_i$.*

The coefficients $\{y_j\}$ and $\{z_k\}$, subject to the constraints in one of the cases of Theorem 6.5, is then a *dual system* L^* to L. If $\{x_i\}$ is feasible for L and attains a value of c for the objective $\sum_i c_i x_i$, and if $\{y_j\}$ and $\{z_k\}$ are feasible for L^* and attain the same c, then they are an *optimal pair*; each half of the pair proves that the other half is optimal. Thus, case 3 of Theorem 6.5 says that every maximization problem in finite linear programming with a finite maximum can be solved by finding an optimal pair.

We cannot directly apply Theorem 6.5 to Problem 6.1 because it is an infinite-dimensional problem. The theorem still holds in infinite dimensions, or in finite dimensions with infinitely many inequalities, with an extra hypothesis such as compactness. We do not know a simple way to make Problem 6.1 compact, but we will find optimal pairs anyway.

The standard notion of an optimal pair from Theorem 6.5 is not exactly the same as that in Section 6.1.2, because Problem 6.1 is nonlinear in the variable A. However, the two concepts are analogous. Indeed, we can change Problem 6.1 slightly to make A a linear variable, as follows. First, we switch to the measure $\hat{\mu} \stackrel{\text{def}}{=} \mu/A$ and divide through by A. Then the first three relations, (35), (36), and (37), are all linear in the variables $\hat{\mu}$ and A. The last two relations, (38) and (39), now have a factor of $1/A$, but we can convert them to these two equations:

$$\int_{\ell,\alpha,\beta} (\omega_{n-1} s_{n,\kappa}^{(-2)}(\ell) - V\ell)\, \mathrm{d}\hat{\mu} \leqslant 0 \qquad \int_{\ell,\alpha,\beta} \ell \, \mathrm{d}\hat{\mu} \geqslant \frac{\omega_{n-1} V}{A}.$$

These relations are moderately weaker than (38) and (39), but switching to them is nearly equivalent to assuming the condition (45). Although the second equation is still nonlinear, it is a convex relation between A and $\hat{\mu}$; it can be expressed by a family of linear relations.

Finally, linear programming in infinite dimensions usually involves topological vector spaces. For instance, the measure μ in Problem 6.1 lies in a space of Borel measures. This raises the question of the appropriate regularity of the dual variable $f(\alpha)$. Because of (35), the function $f(\alpha)$ could in principle be integrable rather continuous; $f(\alpha)\,\mathrm{d}z(\alpha)$

could even be replaced a Borel measure. However, Proposition 6.6 from optimal transport theory tells us that an optimal $f(\alpha)$ is continuous.

6.2.2. *Optimal transport.* We can interpret Problem 6.1 as an optimal transport problem. See Villani [Vil09, Ch.3-5] for background material on optimal transport. Following Villani, we assume that $A\,\mathrm{d}z(\alpha)$ is a distribution of boulangeries and $A\,\mathrm{d}z(\beta)$ is a distribution of cafés. Moreover, for each boulangerie α and café β, there are a range of possible roads parametrized by ℓ. By (35), μ is a transport of baguettes[2] from the boulangeries to the cafés. Problem 6.1 then asks whether the transport is feasible given the constraints that we must pay separate road tolls in Polish zlotys (36), Czech korunas (37), and Hungarian forints (38); and given an exact labor requirement (39) (neither more nor less). Strictly speaking, this is a feasible transport problem rather than an optimal transport problem, but we can convert it to optimal transport.

The function $E(\ell,\alpha,\beta)$ defined in equation (42) is a natural reduction of all four resource limits into one combined cost function, which we can then optimize to test feasibility. In the economics interpretation, the coefficients a, b, and c are currency conversions, while d is a wage rate. The last term $d\ell$ is naturally subtracted if employment is the goal of the program and thus a negative cost. Certainly if any choice of a,b,c,d yields a cost function E such that

$$\int_{\ell,\alpha,\beta} E(\ell,\alpha,\beta)d\mu \leqslant aA^2+2bAV+cV^2-d\omega_{n-1}V \tag{49}$$

is infeasible, then the original multi-resource transport problem is also infeasible. We won't try to prove the converse for all n and κ: that if Problem 6.1 is infeasible, then there exist (a,b,c,d) such that (49) is also infeasible.

Even so, once a,b,c,d are chosen, Problem 6.1 reduces to just (35) and (49), which is a nearly standard optimal transport problem. The two differences are:

1. We have a choice of "roads" parametrized by ℓ. Given a scalar cost, we can convert it to a standard optimal transport problem if we choose the most efficient road for each pair (α,β) and let the cost be
$$E(\alpha,\beta)=\min_{\ell} E(\ell,\alpha,\beta).$$
2. The transport μ does not usually have to be symmetric in α and β. We can live without this constraint because Problem 6.1 is itself symmetric in α and β, if we add the relation $\beta_*(\mu)=A\,\mathrm{d}z(\beta)$, which is the other half of (24). We can symmetrize any solution using
$$\hat{\mu}(\ell,\alpha,\beta)\overset{\text{def}}{=}\frac{\mu(\ell,\alpha,\beta)+\mu(\ell,\beta,\alpha)}{2}.$$

Having fixed a,b,c,d, the remaining dual variable in Problem 6.2 is $f(\alpha)$. Its sole constraint is (43). In optimal transport terminology, $f(\alpha)$ is known as a *Kantorovich potential*. We will call the left side, $F(\ell,\alpha,\beta)$, the *adjusted cost function*. In standard optimal transport, we would have two potentials $f(\alpha)$ and $g(\beta)$ satisfying the equation

$$E(\ell,\alpha,\beta)+f(\alpha)+g(\beta)\geqslant 0. \tag{50}$$

But, just as symmetry is optional in Problem 6.1, it is also optional in Problem 6.2; we can symmetrize a solution to make $f=g$.

[2] Even though in Section 5.2, we transported photons.

Proposition 6.6. *An optimal potential $f(\alpha)$ in Problem 6.2 is a convex function of* $\sec(\alpha)$ *and therefore continuous.*

Proposition 6.6 is a standard type of result in optimal transport theory. A potential that satisfies an equation such as (51) below is called *cost convex*.

Proof. We assume two potentials $f(\alpha)$ and $g(\beta)$. In an asymmetric variation of Problem 6.2, they are chosen to minimize

$$\int_0^{\pi/2} f(\alpha)\,\mathrm{d}z(\alpha) + \int_0^{\pi/2} g(\beta)\,\mathrm{d}z(\beta).$$

For each fixed $g(\beta)$, we can minimize this integral subject to the constraint (50) by choosing

$$f(\alpha) = \sup_{\ell,\beta} \big[-E(\ell,\alpha,\beta) - g(\beta)\big]. \tag{51}$$

For each fixed value of ℓ and β, the supremized function on the right side is linear in $\sec(\alpha)$ by (42). It follows that $f(\alpha)$ is convex in $\sec(\alpha)$ and thus continuous; the same is true of $g(\beta)$. If this asymmetric optimization yields $f \neq g$, then their average $(f+g)/2$ has all of the desired properties. □

7. Proofs of the main results

In this section, we will complete the proofs Theorems 1.4 and 1.5, picking up from Section 6.1. Up to rescaling, we can assume that $\kappa \in \{-1,0,1\}$. We first explicate the condition (47), which we will use to check whether Problem 6.1 has any hope of producing a sharp isoperimetric inequality, and to calculate the parameters in Lemma 6.3.

If $\Omega = B_{n,\kappa}(r)$, then the length of a geodesic chord that makes an angle of α from the normal to $\partial\Omega$ is given by the relation

$$\cos(\alpha) = T_{\kappa,r}(\ell) \stackrel{\text{def}}{=} \begin{cases} \dfrac{\tan(\ell/2)}{\tan(r)} & \text{if } \kappa = 1 \\ \dfrac{\ell}{2r} & \text{if } \kappa = 0. \\ \dfrac{\tanh(\ell/2)}{\tanh(r)} & \text{if } \kappa = -1 \end{cases} \tag{52}$$

Equation (52) thus gives us the function $\ell = h(\alpha)$ in the statement of Lemma 6.4.

We combine equations (42) and (47) to obtain

$$\frac{\partial E}{\partial \ell}(\ell,\alpha,\alpha) = a\frac{s'_{n,\kappa}(\ell)}{\cos(\alpha)^2} + 2b\frac{s_{n,\kappa}(\ell)}{\cos(\alpha)} + cs^{(-1)}_{n,\kappa}(\ell) - d = 0.$$

Combining with (52), we obtain

$$a\frac{s'_{n,\kappa}(\ell)}{T_{\kappa,r}(\ell)^2} + 2b\frac{s_{n,\kappa}(\ell)}{T_{\kappa,r}(\ell)} + cs^{(-1)}_{n,\kappa}(\ell) - d = 0. \tag{53}$$

Again, (53) is an equation for the coefficients a,b,c,d that should hold for $0 \leqslant \ell \leqslant 2r$. In each case, the coefficients will be unique up to rescaling by a positive real number. Note that the factor of $\tan(r)$, r, or $\tanh(r)$ that appears in $T_{\kappa,r}$ factors of out of the question of whether there is a solution, since this factor can be absorbed into the constants a and b.

Our proofs in this section follow a set pattern:

1. Working either from Problem 6.1 or Problem 7.2, and their dual problems, calculate (a,b,c,d) using (53).

2. Change variables from α and β to x and y using (54), (56), or (61). Calculate the cost $E(\ell,x,y)$, the potential $f(x)$, and the adjusted cost $F(\ell,x,y)$ in the new variables.
3. Using calculus methods, establish that the adjusted cost $F(\ell,x,y)$ is non-negative, according to (43). This will fulfill the hypotheses of Lemma 6.3 when $\kappa \geqslant 0$, or its equivalent when $\kappa < 0$, and finish the proof. In the hardest two cases ($n=4$ and $\kappa \neq 0$), this step depends crucially on symbolic algebra software.

Also, we abbreviate $s = s_{n,\kappa}$ throughout.

Here are two general remarks about dimension $n=2$. First, for every value of κ, there is a separation (55) in this dimension. This means that we could have proved the results with a simpler measure $\mu(\ell,\alpha)$ that depends on only one angle, in the spirit of Section 3.2. Second, $a=0$ when $n=2$, so we can immediately accept A as a linear variable in Problem 6.1 or 7.2.

7.1. **Weil's and Croke's theorems.** This case is a warm-up to the more difficult cases with $\kappa \neq 0$. We introduce the change of variables

$$(x,y) \stackrel{\text{def}}{=} \left(\frac{\sec(\alpha)}{r}, \frac{\sec(\beta)}{r}\right) \tag{54}$$

in place of α and β. We will give them the range $x,y \in \mathbb{R}_{\geqslant 0}$. By abuse of notation, we can change variables without changing the names of functions; for example, we can write

$$E(\ell,\alpha,\beta) = E(\ell,\alpha(x),\beta(y)) = E(\ell,x,y).$$

If $\kappa = 0$, then

$$s(\ell) = \ell^{n-1}, \qquad s'(\ell) = (n-1)\ell^{n-2},$$
$$s^{(-1)}(\ell) = \frac{\ell^n}{n}, \qquad s^{(-2)}(\ell) = \frac{\ell^{n+1}}{n(n+1)}.$$

Equation (53) becomes

$$4(n-1)r^2 a\ell^{n-4} + 4rb\ell^{n-2} + \frac{c\ell^n}{n} - d = 0.$$

Obviously this has solutions if $n \in \{2,4\}$ and not otherwise; this point was known to Croke (personal communication).

When $n=2$, the solution is

$$a = 0, \qquad b = \frac{1}{r}, \qquad c = 0, \qquad d = 4.$$

From (42), we thus obtain

$$E(\ell,\alpha,\beta) = \frac{\ell^2(\sec(\alpha)+\sec(\beta))}{2r} - 4\ell$$

Then (46) and (52) give us the potential

$$f(\alpha) = -\frac{E(2r\cos(\alpha),\alpha,\alpha)}{2} = 2r\cos(\alpha).$$

Then the adjusted cost (43) separates as

$$F(\ell,\alpha,\beta) = G(\ell,\alpha) + G(\ell,\beta) \tag{55}$$

with

$$G(\ell,\alpha) = \frac{\ell^2\sec(\alpha)}{2r} - 2\ell + 2r\cos(\alpha) = \frac{2r\cos(\alpha)-\ell}{2r\cos(\alpha)} \geqslant 0.$$

Thus $F \geqslant 0$, which establishes Weil's theorem.

When $n = 4$, the solution to (53) is

$$a = \frac{1}{r^2}, \qquad b = 0, \qquad c = 0, \qquad d = 12.$$

These coefficients plainly satisfy condition (45). The cost function is

$$E(\ell,\alpha,\beta) = \frac{\ell^3 \sec(\alpha)\sec(\beta)}{r^2} - 12\ell,$$

the potential is

$$f(\alpha) = -\frac{E(2r\cos(\alpha),\alpha,\alpha)}{2} = 8r\cos(\alpha),$$

and their sum is

$$F(\ell,\alpha,\beta) = \frac{\ell^3 \sec(\alpha)\sec(\beta)}{r^2} - 12\ell + 8r(\cos(\alpha)+\cos(\beta)).$$

Using the change of variables (54),

$$F(\ell,x,y) = \ell^3 xy - 12\ell + \frac{8}{x} + \frac{8}{y}.$$

We want to show that $F \geqslant 0$. For each fixed value of xy, F is minimized when $x = y$. We can then calculate

$$F(\ell,x,x) = \ell^3 x^2 - 12\ell + \frac{16}{x} = \frac{(\ell x+4)(\ell x-2)^2}{x} \geqslant 0.$$

This establishes Croke's theorem.

Following the comments after the proof of Theorem 1.6 in Section 3.2, our proof of Croke's theorem is only superficially different from Croke's proof. The extra point here is that Croke's theorem (and Weil's theorem along with it) hold in Model 6.1, which establishes part of Theorem 1.17.

7.2. **The positive case.** In this section we will establish Theorem 1.4. We will let $\kappa = 1$, but before we do that, we note that $\kappa = 0$ is a limiting case of $\kappa > 0$. Section 7.1 established that a sharp result in the case $\kappa = 0$ is only possible when $n \in \{2,4\}$, this justifies the same restriction in Theorem 1.4.

We use the change of variables

$$(x,y) \stackrel{\text{def}}{=} \Big(\frac{\sec(\alpha)}{\tan(r)}, \frac{\sec(\beta)}{\tan(r)}\Big) \tag{56}$$

with the range $x,y \in \mathbb{R}_{\geqslant 0}$. Note that equation (52) simplifies to

$$\tan(\frac{\ell}{2}) = \frac{1}{x}. \tag{57}$$

7.2.1. *Dimension 2.* In dimension $n = 2$,

$$s(\ell) = \sin(\ell), \qquad s'(\ell) = \cos(\ell),$$
$$s^{(-1)}(\ell) = 1-\cos(\ell), \qquad s^{(-2)}(\ell) = \ell - \sin(\ell)$$

when $\ell < \pi$, and

$$s(\ell) = 0, \qquad s^{(-1)}(\ell) = 2, \qquad s^{(-2)} = 2\ell - \pi$$

for $\ell \geqslant \pi$. Equation (53), with (52), becomes

$$\frac{a\tan(r)^2\cos(\ell)}{\tan(\ell/2)^2} + \frac{2b\tan(r)\sin(\ell)}{\tan(\ell/2)} + c(1-\cos(\ell)) - d = 0.$$

The solution is

$$a = 0, \qquad b = \frac{1}{\tan(r)}, \qquad c = 2, \qquad d = 4.$$

In the variables (56), the cost function (42) is

$$E(\ell,x,y) = (1-\cos(\ell))(x+y) - 2\sin(\ell) - 2\ell,$$

for $\ell \leqslant \pi$, and is constant in ℓ for $\ell \geqslant \pi$:

$$E(\ell,x,y) = E(\pi,x,y) \qquad \forall \ell \geqslant \pi. \tag{58}$$

Using (57), the potential (46) becomes

$$f(x) = \ell = 2\arctan(\frac{1}{x}).$$

The adjusted cost (43) again separates according to (55), where this time

$$G(\ell,x) = (1-\cos(\ell))x - \sin(\ell) - \ell + 2\arctan(\frac{1}{x}).$$

We can minimize G with the derivative test either in ℓ or in x. The latter is slightly simpler and gives us

$$\frac{\partial G}{\partial x}(\ell,x) = \frac{x^2(1-\cos(\ell)) - (\cos(\ell)+1)}{x^2+1}.$$

We learn that $\partial G/\partial x$ crosses 0 exactly once, when x and ℓ satisfy (57); this is therefore the minimum of G for each fixed ℓ. Since the relation (57) is used to define the potential $f(x)$, it is automatic that this minimum value is 0; the substitution $x = 1/\tan(\ell/2)$ also establishes it. Thus $G(\ell,x) \geqslant 0$, which confirms (43) and establishes the $n = 2$ case of Theorem 1.4.

7.2.2. *Dimension 4.* In dimension $n = 4$,

$$\begin{aligned} s(\ell) &= \sin(\ell)^3, \\ s'(\ell) &= 3\sin(\ell)^2\cos(\ell), \\ s^{(-1)}(\ell) &= \frac{\cos(\ell)^3 - 3\cos(\ell) + 2}{3}, \\ s^{(-2)}(\ell) &= \frac{6\ell - \sin(\ell)^3 - 6\sin(\ell)}{9} \end{aligned}$$

when $\ell < \pi$, and

$$s(\ell) = 0, \quad s^{(-1)}(\ell) = \frac{4}{3}, \quad s^{(-2)}(\ell) = \frac{4\ell - 2\pi}{3}$$

when $\ell \geqslant \pi$. Equation (53) becomes

$$\frac{3a\tan(r)^2\cos(\ell)\sin(\ell)^2}{\tan(\ell/2)^2} + \frac{2b\tan(r)\sin(\ell)^3}{\tan(\ell/2)} + \frac{c(\cos(\ell)^3 - 3\cos(\ell) + 2)}{3} - d = 0.$$

The solution is

$$a = \frac{1}{\tan(r)^2}, \qquad b = \frac{3}{\tan(r)}, \qquad c = 9, \qquad d = 12.$$

The clean optimality condition (45) becomes

$$9V < 12\omega_3 = 24\pi^2.$$

Since V is at most the volume of a hemisphere, we have

$$9V < \frac{9\omega_4}{2} = 12\pi^2.$$

Thus (45) holds.

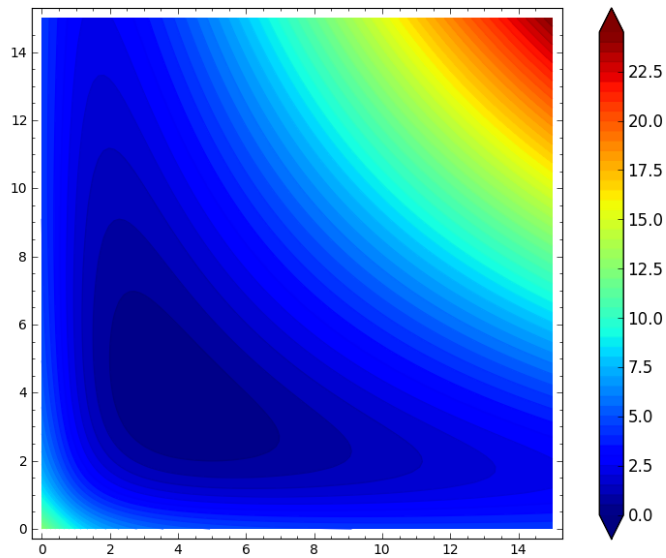

FIGURE 8. The slice $F(\pi/6, x, y)$ in the case $\kappa = 1$.

The cost function (42) is

$$E(\ell, x, y) = \sin(\ell)^3 xy + (\cos(\ell)^3 - 3\cos(\ell) + 2)(x+y) - \sin(\ell)^3 - 6\sin(\ell) - 6\ell.$$

for $\ell \leqslant \pi$, while once again E is constant in ℓ for $\ell \geqslant \pi$, as in (58). The potential from (46) and (57) is

$$f(x) = 6\arctan(\frac{1}{x}) + \frac{2x}{x^2+1}.$$

We will include the values $x = 0$ and $y = 0$ in our calculations, so it is helpful to recall that

$$\arctan(\frac{1}{x}) = \frac{\pi}{2} - \arctan(x).$$

The adjusted cost (43) is

$$\begin{aligned} F(\ell, x, y) = \sin(\ell)^3 xy + (\cos(\ell)^3 - 3\cos(\ell) + 2)(x+y) - \sin(\ell)^3 - 6\sin(\ell) - 6\ell \\ + 6\pi - 6\arctan(x) + \frac{2x}{x^2+1} - 6\arctan(y) + \frac{2y}{y^2+1}. \end{aligned} \tag{59}$$

The remainder of the proof of Theorem 1.4 is given by the following lemma. Although the lemma is evident from contour plots (*e.g.*, Figure 8), the authors found it surprisingly tricky to find a rigorous proof.

Lemma 7.1. *The function $F(\ell, x, y)$ on $[0, \pi] \times \mathbb{R}^2_{\geqslant 0}$ given by (59) is non-negative, and vanishes only when*

$$x = y = \frac{1}{\tan(\ell/2)}.$$

Proof. We will use these immediate properties of the potential $f(x)$:

$$f(0) = 3\pi, \qquad f(x) > 0.$$

We first check the non-compact direction of the domain of F. There exists a constant $k > 0$ such that

$$s^{(-1)}(\ell) \geqslant k\ell^4.$$

(Because $\ell^4/s^{(-1)}(\ell)$ is continuous on $[0,\pi]$ and therefore bounded. In fact

$$k = \frac{s^{(-1)}(\pi)}{\pi^4} = \frac{4}{3\pi^4}$$

works.) Thus

$$\begin{aligned} F(\ell,x,y) &= s(\ell)xy + 3s^{(-1)}(\ell)(x+y) + 9s^{(-2)}(\ell) - 12\ell + f(x) + f(y) \\ &\geqslant 3k(x+y)\ell^4 - 12\ell \end{aligned}$$

by discarding positive terms and simplifying $s^{(-1)}(\ell)$. Thus

$$\begin{aligned} \liminf_{x+y\to\infty} \big(\min_{\ell} F(\ell,x,y)\big) &\geqslant \liminf_{x+y\to\infty} \Big(\min_{\ell\geqslant 0} \big(3k(x+y)\ell^4 - 12\ell\big)\Big) \\ &= \liminf_{x+y\to\infty} \frac{-9}{\sqrt[3]{k(x+y)}} = 0. \end{aligned}$$

The inequality comes from discarding positive terms, while the equality follows just from the properties of $s^{(-1)}(\ell)$ that it is continuous, and that it is positive for $\ell > 0$.

Having confined the locus of $F(\ell,x,y) \leqslant -\varepsilon$ to a compact region for every $\varepsilon > 0$, we will calculate derivatives and boundary values to show that this region cannot have a local minimum and must thus be empty. First, taking $\ell = 0$, we get

$$F(0,x,y) = f(x) + f(y) > 0.$$

Second, taking $\ell = \pi$, we get

$$F(\pi,x,y) = 4(x+y) - 6\pi + f(x) + f(y).$$

Here we check that

$$\frac{\partial F}{\partial x}(\pi,x,y) = \frac{4x^4}{(x^2+1)^2} \geqslant 0, \qquad \frac{\partial F}{\partial y}(\pi,x,y) = \frac{4y^4}{(y^2+1)^2} \geqslant 0, \qquad F(\pi,0,0) = 0.$$

Fourth, taking $x = y = 0$, we obtain

$$F(\ell,0,0) = 9s^{(-2)}(\ell) - 12\ell + 6\pi.$$

We check in this case that

$$F(\pi,0,0) = 0, \qquad \frac{\partial F}{\partial \ell}(\ell,0,0) = 9s^{(-1)}(\ell) - 12 \leqslant 0.$$

The fifth case is the case $y = 0$ with x and ℓ interior, which by symmetry is equivalent to the case $x = 0$ with y and ℓ interior. The sixth and final case is the interior for all three coordinates. We will handle the fifth and sixth cases together. Using the final change of variables

$$t \stackrel{\text{def}}{=} \tan\big(\frac{\ell}{2}\big),$$

and with the help of Sage, we learn that

$$\frac{\partial F}{\partial \ell}(t,x,y) = -12\frac{(t^4-t^2)xy-2t^3(x+y)+3t^2+1}{(t^2+1)^3}$$
$$\frac{\partial F}{\partial x}(t,x,y) = 4\frac{(2t^3y-3t^2-1)(x^2+1)^2+x^4(t^2+1)^3}{(t^2+1)^3(x^2+1)^2}.$$

Note that the partial derivatives are rational functions in x, y, and t. We can rigorously determine the common zeroes of their numerators by finding their associated prime ideals in the ring $\mathbb{Q}[x,y,t]$ using the "`associated_primes`" function in Sage[3] (In other words, we use the Lasker-Noether factorization theorem converted to an algorithm by the Gröebner basis method.) The solution set is characterized by five prime ideals:

$$\begin{aligned}
I_1 &= (x-y, yt-1)\\
I_2 &= (x+t, t^2+1)\\
I_3 &= (y+t, t^2+1)\\
I_4 &= (x, 2yt^3-3t^2-1)\\
I_5 &= (2x^2y+3x^2t+xyt+x+y, x^2t^3+xyt^3-xt^2-yt^2-2x+2t,\\
&\qquad xy^2t^2+2xyt^3+y^2t^3-2xyt-3xt^2-yt^2-y-t,\\
&\qquad y^2t^4+y^2t^2+xt^3+3yt^3+2xy+3xt-3yt-7t^2-1,\\
&\qquad xyt^4-xyt^2-2xt^3-2yt^3+3t^2+1).
\end{aligned}$$

Four of these ideals cannot vanish when $x>0$ and $y,\ell \geqslant 0$: I_2 and I_3 contain t^2+1, I_4 contains x, and I_5 contains

$$2x^2y+3x^2t+xyt+x+y > 0.$$

The ideal I_1 yields the desired locus $x = y = 1/t$.

A careful examination of the equality cases shows that $x = y = 1/t$ is the only possibility for the minimum value $F = 0$. □

7.3. **The negative case.** In this section we will establish Theorem 1.5. As in Section 7.2, we let $\kappa = -1$ and we must take $n \in \{2,4\}$. We cannot use Problem 6.1, because in both dimensions, one of the dual coefficients turns out to be negative. Instead we the use the following model, which is provided by Lemma 5.5.

LP Problem 7.2. *Given n, A, V, and L, let*

$$q = \frac{n-1}{\tanh(L)}.$$

[3]See the attached Sage files in the source file of the arXiv version of this paper.

Is there a symmetric, positive measure $\mu(\ell,\alpha,\beta)$ *such that*

$$\alpha_*(\mu) = \int_{\ell,\beta} \mathrm{d}\mu = A\,\mathrm{d}z(\alpha)$$
$$\int_{\ell,\alpha,\beta} \Big(s(\ell)\sec(\beta) - qs^{(-1)}(\ell)\Big)\sec(\alpha)\,\mathrm{d}\mu_\Omega \leqslant A^2 - qAV.$$
$$\int_{\ell,\alpha,\beta} \big(s^{(-1)}(\ell)\sec(\alpha) - qs^{(-2)}(\ell)\big)\,\mathrm{d}\mu_\Omega \leqslant AV - qV^2$$
$$\int_{\ell,\alpha,\beta} s^{(-2)}(\ell)\,\mathrm{d}\mu \leqslant V^2$$
$$\int_{\ell,\alpha,\beta} \ell\,\mathrm{d}\mu = \omega_{n-1}V?$$

We will need the dual problem, which we can state without changing variables.

LP Problem 7.3 (Dual to Problem 7.2)**.** *Given n, A, V, and L, let*

$$q = \frac{n-1}{\tanh(L)}.$$

Are there numbers $a,b,c,d \in \mathbb{R}$ *and a continuous function* $f : [0,\pi/2) \to \mathbb{R}$ *such that*

$$a \geqslant 0 \qquad 2b+qa \geqslant 0 \qquad c+q(2b+qa) \geqslant 0 \tag{60}$$
$$\begin{aligned} &as_{n,-1}(\ell)\sec(\alpha)\sec(\beta) + bs^{(-1)}_{n,-1}(\ell)\big(\sec(\alpha)+\sec(\beta)\big) \\ &\qquad + cs^{(-2)}_{n,-1}(\ell) - d\ell + f(\alpha) + f(\beta) \geqslant 0 \end{aligned}$$
$$aA^2 + 2bAV + cV^2 - d\omega_{n-1}V + 2A\int_0^{\pi/2} f(\alpha)\,\mathrm{d}z(\alpha) < 0?$$

We will use the change of variables

$$(x,y) \stackrel{\text{def}}{=} \Big(\frac{\sec(\alpha)}{\tanh(r)}, \frac{\sec(\beta)}{\tanh(r)}\Big) \tag{61}$$

with the range $x,y \in (1,\infty)$. Equation (52) simplifies to

$$\tanh(\frac{\ell}{2}) = \frac{1}{x}. \tag{62}$$

7.3.1. *Dimension* 2. In dimension $n=2$,

$$s(\ell) = \sinh(\ell), \qquad s'(\ell) = \cosh(\ell),$$
$$s^{(-1)}(\ell) = \cosh(\ell) - 1, \qquad s^{(-2)}(\ell) = \sinh(\ell) - \ell.$$

Equation (53), with (52), becomes

$$\frac{a\tanh(r)^2\cosh(\ell)}{\tanh(\ell/2)^2} + \frac{2b\tanh(r)\sinh(\ell)}{\tanh(\ell/2)} + c(\cosh(\ell)-1) - d = 0.$$

The solution is

$$a = 0, \qquad b = \frac{1}{\tanh(r)}, \qquad c = -2, \qquad d = 4.$$

We need to check the third case of condition (60), which reduces to

$$c + 2qb = -2 + \frac{2}{\tanh(r)\tanh(L)} \geqslant 0.$$

Since the tanh function is bounded above by 1, this is immediate.

In the variables (61), the cost function (42) is

$$E(\ell,x,y) = (\cosh(\ell)-1)(x+y) - \sinh(\ell) - 2\ell.$$

The potential (46) is

$$f(x) = 2\operatorname{arctanh}(\frac{1}{x}).$$

The adjusted cost (43) separates according to (55) with

$$G(\ell,x) = (\cosh(\ell)-1)x - \sinh(\ell) - \ell + 2\operatorname{arctanh}(\frac{1}{x}).$$

We minimize G using the derivative test in x to obtain

$$\frac{\partial G}{\partial x}(\ell,x) = \frac{x^2(\cosh(\ell)-1) - (\cosh(\ell)+1)}{x^2-1}.$$

We learn that the minimum of G in x for each fixed ℓ occurs when x and ℓ satisfy (62) and it is easy to confirm that the value is 0. Thus $G(\ell,x) \geqslant 0$, which confirms (43) and establishes the $n=2$ case of Theorem 1.5.

7.3.2. *Dimension* 4. In dimension $n=4$,

$$\begin{aligned} s(\ell) &= \sinh(\ell)^3, \\ s'(\ell) &= 3\cosh(\ell)\sinh(\ell)^2, \\ s^{(-1)}(\ell) &= \frac{\cosh(\ell)^3 - 3\cosh(\ell)+2}{3}, \\ s^{(-2)}(\ell) &= \frac{\sinh(\ell)^3 - 6\sinh(\ell)+6\ell}{9}. \end{aligned}$$

Equation (53) becomes

$$\frac{3a\tanh(r)^2\cosh(\ell)\sinh(\ell)^2}{\tanh(\ell/2)^2} + \frac{2b\tanh(r)\sinh(\ell)^3}{\tanh(\ell/2)} + \frac{c(\cosh(\ell)^3-3\cosh(\ell)+2)}{3} - d = 0.$$

The solution is

$$a = \frac{1}{\tanh(r)^2}, \qquad b = -\frac{3}{\tanh(r)}, \qquad c = 9, \qquad d = 12.$$

We need to check the second case of condition (60):

$$2b+qa = -\frac{6}{\tanh(r)} + \frac{3}{\tanh(L)\tanh(r)^2} \geqslant 0.$$

This condition is equivalent to the smallness hypothesis (4).

Following the notation of Lemma 6.3, we claim that our choices for a,b,c,d and a suitable choice of f produce a polynomial $P(A)$ with two real roots $A_0 < A_1$, where A_1 is the sharp isoperimetric value, that together are a solution to Problem 7.3 when $A \in (A_0,A_1)$. In other words, we want to confirm (43), but this is the hardest part of the proof and we save it for the end of the section. In the meantime, we settle a different and simpler difficulty. Recall that

$$V = |B_{4,-1}(r)| = \omega_3 s^{(-1)}(r), \qquad A_1 = |\partial B_{4,-1}(r)| = \omega_3 s(r),$$

and that $A_0 < 0$ is equivalent to the clean optimality condition (45). Since $s^{(-1)}(r)$ is unbounded, (45) does not hold for all V. This criterion is not needed to prove Theorem 1.5, since the set of geometrically feasible A is the closed ray $[A_1,\infty)$; but it is important for Theorem 1.17, which asserts that every $A \in [0,A_1)$ is infeasible for Problem 7.2. To prove this part of Theorem 1.17, we claim a second set of values a_2, b_2, c_2, d_2, f_2 that are feasible for Problem 7.3 and that produce $P_2(A)$ with $P_2(0) < 0$ and $P_2(A_1) < 0$.

Let $A_2 = 3\tanh(r)V$. On the one hand, the proof of Theorem 1.16 in Section 8.5.1 produces the feasible values

$$a_2 = 0, \qquad b_2 = 1, \qquad c_2 = -3, \qquad d_2 = 0, \qquad f_2 = 0,$$

and a linear $P_2(A)$ that vanishes at A_2. Thus, every $A \in [0, A_2)$ is infeasible. On the other hand, we claim that $P(A_2) < 0$. We use $P(A_1) = 0$ to calculate that

$$\int_0^{\pi/2} f(\alpha)\,\mathrm{d}z(\alpha) = \frac{d\omega_3 V - aA_1^2 - 2bA_1V - cV^2}{2A_1} = \frac{4\pi^2(\cosh(r)-1)}{\sinh(r)}$$

and

$$P(A_2) = \Big(\frac{A_2}{\tanh(r)} - 3V\Big)^2 - 12\omega_3 V + 2A_2\int_0^{\pi/2} f(\alpha)\,\mathrm{d}z(\alpha) = -\frac{24\pi^2}{\cosh(r)} < 0,$$

so that $A_2 \in (A_0, A_1)$, as desired.

It remains to confirm (43) for the given values of a, b, c, d. The cost function (42) is

$$E(\ell,x,y) = \sinh(\ell)^3 xy - (\cosh(\ell)^3 - 3\cosh(\ell) + 2)(x+y) + \sinh(\ell)^3 - 6\sinh(\ell) - 6\ell.$$

The potential from (46) and (62) is

$$f(x) = 6\operatorname{arctanh}(\frac{1}{x}) + \frac{2x}{x^2-1}.$$

The adjusted cost (43) is

$$\begin{aligned} F(\ell,x,y) = {} & \sinh(\ell)^3 xy - (\cosh(\ell)^3 - 3\cosh(\ell) + 2)(x+y) + \sinh(\ell)^3 - 6\sinh(\ell) - 6\ell \\ & + 6\operatorname{arctanh}(\frac{1}{x}) + \frac{2x}{x^2-1} + 6\operatorname{arctanh}(\frac{1}{y}) + \frac{2y}{y^2-1}. \end{aligned} \tag{63}$$

We conclude the proof of Theorem 1.5 with the following lemma. The lemma is also numerically evident but surprisingly tricky (for the authors).

Lemma 7.4. *The function $F(\ell,x,y)$ on $\mathbb{R}_{\geqslant 0} \times (1,\infty)^2$ given by (63) is non-negative, and vanishes only when*

$$x = y = \frac{1}{\tanh(\ell/2)}.$$

Proof. The proof is analogous to that of Lemma 7.1, but differs in its technical details. Throughout the proof, we will fix y and minimize $F(\ell,x,y)$ with respect to x and ℓ.

To check the non-compact limits of x and ℓ, we re-express F as:

$$F(\ell,x,y) = \sinh(\ell)^3(x-1)(y-1) + h(\ell)(x+y) - 6\sinh(\ell) - 6\ell + f(x) + f(y),$$

where

$$h(\ell) = (\sinh(\ell)^3 - \cosh(\ell)^3 + 3\cosh(\ell) - 2) = \frac{(3e^\ell + 1)(1-e^{-\ell})^3}{4} > 0.$$

We also have

$$f(x) = \frac{1}{x-1} + \frac{1}{x+1} + \operatorname{arctanh}(\frac{1}{x}) > \frac{1}{x-1}$$

and the elementary relation $\sinh(\ell) \geqslant \ell$. We combine these comparisons to obtain the bound

$$\hat{F}(\ell,x,y) \stackrel{\text{def}}{=} \sinh(\ell)^3(x-1)(y-1) + \frac{1}{x-1} - 12\sinh(\ell) < F(\ell,x,y).$$

The function $\hat{F}$ is useful for minimizing with respect to either ℓ or x, leaving the other variables fixed. It is a bit simpler to use the variables

$$(x_1, y_1) \stackrel{\text{def}}{=} (x-1, y-1),$$

which we will need anyway later in the proof. We obtain

$$\begin{aligned}
\hat{F}(\ell, x_1, y_1) &= \sinh(\ell)^3 x_1 y_1 + \frac{1}{x_1} - 12\sinh(\ell) \\
\min_{\ell} \hat{F}(\ell, x_1, y_1) &= \frac{-16}{\sqrt{x_1 y_1}} + \frac{1}{x_1} \\
\min_{x_1} \hat{F}(\ell, x_1, y_1) &= 2\sqrt{\sinh(\ell)^3 y_1} - 12\sinh(\ell).
\end{aligned}$$

We obtain these uniform lim infs:

$$\begin{aligned}
\liminf_{x\to\infty} \big(\inf_{\ell} F(\ell,x,y)\big) &\geqslant \lim_{x_1\to\infty} \big(\min_{\ell} \hat{F}(\ell,x_1,y_1)\big) = 0 \\
\liminf_{x\to 1} \big(\inf_{\ell} F(\ell,x,y)\big) &\geqslant \lim_{x_1\to 0} \big(\min_{\ell} \hat{F}(\ell,x_1,y_1)\big) = \infty \\
\liminf_{\ell\to\infty} \big(\inf_{x} F(\ell,x,y)\big) &\geqslant \lim_{\ell\to\infty} \big(\min_{x_1} \hat{F}(\ell,x_1,y_1)\big) = \infty \\
\liminf_{\ell\to 0} \big(\inf_{x} F(\ell,x,y)\big) &\geqslant \lim_{\ell\to 0} \big(\min_{x_1} \hat{F}(\ell,x_1,y_1)\big) = 0.
\end{aligned}$$

Once we control x, we can also check the last case more directly by calculating that

$$F(0,x,y) = f(x) + f(y) > 0.$$

Either way, this establishes that we can use the derivative test for each fixed y to confirm that $F(\ell,x,y) \geqslant 0$.

We use the final change of variables

$$t \stackrel{\text{def}}{=} \tanh(\frac{\ell}{2}).$$

Sage tells us that

$$\begin{aligned}
\frac{\partial F}{\partial \ell}(t,x,y) &= -12\frac{(t^4+t^2)xy - 2t^3(x+y) + 3t^2 - 1}{(t^2-1)^3}, \\
\frac{\partial F}{\partial x}(t,x,y) &= -4\frac{(2t^3y - 3t^2 + 1)(x^2-1)^2 + x^4(t^2-1)^3}{(t^2-1)^3(x^2-1)^2}.
\end{aligned}$$

We again rigorously determine the common zeroes of their numerators by finding their associated prime ideals in the ring $\mathbb{Q}[x,y,t]$ using Sage. The solution set in this case is characterized by 7 prime ideals:

$$
\begin{aligned}
I_1 &= (x-y, yt-1)\\
I_2 &= (t+1, x+1)\\
I_3 &= (t+1, y+1)\\
I_4 &= (t-1, x-1)\\
I_5 &= (t-1, y-1)\\
I_6 &= (x, 2yt^3-3t^2+1)\\
I_7 &= (2x^2y-3x^2t-xyt-x-y, x^2t^3+xyt^3-xt^2-yt^2+2x+2t,\\
&\qquad xy^2t^2-2xyt^3-y^2t^3-2xyt+3xt^2+yt^2-y+t,\\
&\qquad y^2t^4-y^2t^2+xt^3+3yt^3+2xy-3xt+3yt-7t^2+1,\\
&\qquad\qquad xyt^4+xyt^2-2xt^3-2yt^3+3t^2-1).
\end{aligned}
$$

The ideal I_1 yields the desired locus $x=y=1/t$, while the other six do not vanish when $0 \leqslant t < 1$ and $x,y > 1$. Five of these cases are easy: The ideals I_2 and I_3 contain $t+1$, the ideals I_4 and I_5 contain $t-1$, and the ideal I_6 contains x.

The ideal I_7 is obviously more complicated. Setting the first two generators to zero, we obtain:

$$
\begin{aligned}
2x^2y-3x^2t-xyt-x-y &= 0\\
x^2t^3+xyt^3-xt^2-yt^2+2x+2t &= 0.
\end{aligned}
$$

Since the first equation is linear in t, we can eliminate it by substitution, and then clear the denominator and eliminate a factor of $x+y$. The resulting equation in x and y is

$$4x^6y^3-12x^5y^2-8x^4y^3+27x^5+27x^4y+17x^3y^2+5x^2y^3-11x^3-11x^2y-5xy^2-y^3=0.$$

We can express this in the variables x_1 and y_1 as

$$
\begin{aligned}
&4x_1^6y_1^3+12x_1^6y_1^2+24x_1^5y_1^3+12x_1^6y_1+60x_1^5y_1^2+52x_1^4y_1^3+4x_1^6+48x_1^5y_1+96x_1^4y_1^2\\
&\quad+48x_1^3y_1^3+39x_1^5+63x_1^4y_1+41x_1^3y_1^2+8x_1^2y_1^3+154x_1^4+46x_1^3y_1+312x_1^3+55x_1^2y_1\\
&\qquad+336x_1^2+56x_1y_1+176x_1+16y_1+32+9x_1^2y_1(y_1-1)^2+2x_1y_1(y_1-2)^2=0.
\end{aligned}
$$

The left side is manifestly a sum of positive terms when $x_1,y_1 > 0$ and thus cannot vanish. Thus I_7 cannot vanish when $x,y > 1$, which completes the derivative test for $F(\ell,x,y)$. □

8. Proofs of other results

8.1. **Uniqueness.** Problems 6.1 and 7.2 both place strong restrictions on μ and therefore on Ω in the sharp case. First, all of the inequalities in Problem 6.1 become equalities when $\kappa > 0$; all of the inequalities in Problem 7.2 become equalities when $\kappa < 0$. In particular, equation (27) becomes an equality, which implies that Ω is convex and that the candle comparison is an equality at short distances. That in turn implies that Ω satisfies $\mathrm{Ric} \geqslant (n-1)\kappa g$ and that it is the equality case of Bishop's inequality [BC64, Sec. 11.10], which implies that it has constant curvature $K=\kappa$.

The case $\kappa = 0$ does not use (27), but it does use (25). This again implies that Ω is convex. The stronger assumption that Ω is $\sqrt{\mathrm{Ric}}$ class 0 together with equality in (25) tells us again that Ω has constant curvature $K=0$.

Second, sharpness tells us that μ_Ω is concentrated on the locus given by equation (52). In other words, every chord in Ω has the same length and incident angles as if Ω were a round ball $B_{n,\kappa}(r)$. If Ω is convex with constant curvature, this implies that Ω is isometric to $B_{n,\kappa}(r)$.

8.2. **Günther's inequality with reflections.** In this section, we will prove Proposition 5.8.

If $\gamma(t)$ is a smooth curve in M with $t \in [0,r]$, then it is a constant-speed geodesic if and only if it is a critical point of the energy functional

$$E(\gamma) = \int_0^r \frac{\langle \gamma'(t), \gamma'(t)\rangle}{2}\, \mathrm{d}t$$

assuming Dirichlet boundary conditions (*i.e.*, that we fixed the endpoints of γ). Let γ be such a geodesic with unit speed, and let $y(t)$ be a smooth, infinitesimal normal displacement. Then we can define a relative energy

$$E(y) \stackrel{\text{def}}{=} E(\gamma+y) - E(\gamma) + O(||y||^3),$$

which is just the second variational derivative of the curve energy, equivalently half of the second variation of the curve length. We can identify the normal bundle to $\gamma(t)$ with $\mathbb{R}^{n-1}$ using parallel transport, thus view y as a function with values $y(t) \in \mathbb{R}^{n-1}$. If γ is an ordinary geodesic without reflections, then by a standard calculation,

$$E(y) = \int_0^r \left[\langle y'(t), y'(t)\rangle - \langle y(t), R(t)y(t)\rangle\right] \mathrm{d}t,$$

where

$$R(t) = R(\cdot, \gamma'(t), \cdot, \gamma'(t))$$

is the Riemann curvature tensor specialized at the unit tangent γ'. This leads to the differential equation

$$y''(t) = -R(t)y(t), \tag{64}$$

which is satisfied by y when it is a Jacobi field, *i.e.*, a geodesic displacement of γ.

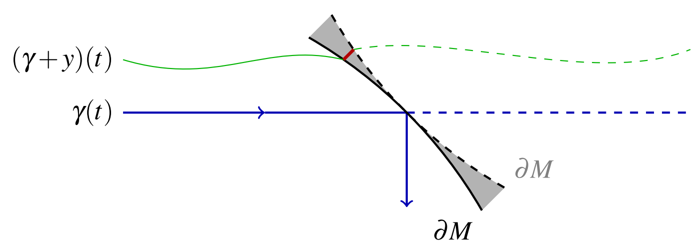

FIGURE 9. Diagram of a vector field y that displaces a geodesic γ in a curved surface (non-geodesically), and a continuation if γ were straight. The short red segment is the length variation of $\gamma+y$ due to the reflection.

If γ reflects from ∂M, then the energy has extra terms. We will derive the energy (65) and a modified Jacobi field equation (66). Although these equation are not really new [Inn98, Sec. 2], we give a geometric argument that we have not seen elsewhere. To understand the extra terms in $E(y)$ due to the reflections, suppose that γ reflects from ∂M at a point $p = \gamma(t)$, and let $Q = Q(p)$ be shape operator ∂M relative to the inward unit normal $w = w(p)$, *i.e.*,

$$Qu = -\nabla_u w.$$

If we give γ a ghost extension as in Figure 9, then the displacement $\gamma+y$ has a gap when ∂M is curved. (The figure shows the convex case with a positive gap; the gap can also have negative width.) We first assume the simplest case in which γ is normal to ∂M. The quadratic form $\langle\cdot,Q\cdot\rangle/2$ osculates ∂M, so that the width of the gap, and thus the negative of the change in length, is $\langle y,Qy\rangle$. If the angle of incidence of γ is $\theta \neq 0$, then this answer is subject to two corrections. First, the gap is at an angle of θ from γ, so the length saved is $\cos(\theta)\langle\cdot,Q\cdot\rangle$. Second, y no longer represents the position that $\gamma+y$ meets $T_p\partial M$, again because the surface is angled.

To derive where $\gamma+y$ meets $T_p\partial M$, we call $T_p(\partial M)$ the *tangent hyperplane*, the normal $N_p(\gamma)$ to $\gamma(t)$ the *coronal hyperplane*, and the 2-dimensional plane spanned by $w(p)$ and $\gamma'(t)$ the *sagittal plane*[4]. Let P be the orthogonal projection from the tangent hyperplane to the coronal hyperplane. If we choose an orthonormal coronal basis $e_1,\dots,e_{n-1}$ such that e_1 is in the sagittal plane, and a matching tangent basis, then

$$P = \begin{pmatrix} \cos(\theta) & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Then the change in length, and therefore the extra energy term, is

$$-\cos(\theta)\langle P^{-1}y, QP^{-1}y\rangle.$$

(This formula still works when $\theta = 0$ if we take P to be the identity matrix.) If γ reflects from a sequence of boundary points $\{p_k\}$ at times $\{t_k\}$, then we have the same change in length using angles θ_k and symmetric matrices P_k and Q_k, and we can abbreviate the result by letting

$$A_k \stackrel{\text{def}}{=} \cos(\theta_k)P_k^{-1}Q_kP_k^{-1}.$$

Then energy of the normal field y is

$$E(y) = \int_0^r \left[\langle y'(t),y'(t)\rangle - \langle y(t),R(t)y(t)\rangle\right] \mathrm{d}t - \sum_k \langle y(t_k),A_k y(t_k)\rangle. \tag{65}$$

Thus, if y is a (reflecting) Jacobi field, it satisfies the distributional differential equation

$$y''(t) = -R(t)y(t) - \sum_k A_k y(t_k)\delta_{t_k}(t), \tag{66}$$

where δ_t is a Dirac delta measure on $\mathbb{R}$ concentrated at t. Note that if ∂M is concave, then Q_k is negative semidefinite and therefore so is A_k.

We now follow a standard proof of Günther's inequality [GHL90, Thm. 3.101]. First, we will need that the energy (65) is positive definite, so that if y Jacobi field, it is an energy minimum (assuming Dirichlet boundary conditions) and not just a critical point. This is standard in the proof of Günther's inequality without the A_k terms, with the aid of the length restriction when $\kappa > 0$. It is still true with the A_k terms, since each such term is positive semidefinite.

Second, we consider a matrix solution Y to (66) with $Y(0) = 0$ and $Y(r) = I$. Then the candle function of γ satisfies

$$j(\gamma,0,\ell) = \frac{\det Y(\ell)}{\det Y(r)},$$

[4]This terminology is borrowed from human anatomy.

and the logarithmic derivative at r is given by

$$\frac{\partial}{\partial t}\Big|_{t=r} \log(j(\gamma,0,t)) = (\det Y)'(r) = \mathrm{Tr}(Y)(r).$$

We generalize the energy (65) to the matrix argument Y, and we interpret it as a function of Y, $R(t)$, and each A_k:

$$E(Y,R,A) = \int_0^r \big[\langle Y'(t),Y'(t)\rangle - \langle Y(t),R(t)Y(t)\rangle\big]\,\mathrm{d}t - \sum_k \langle Y(t_k),A_k Y(t_k)\rangle, \qquad (67)$$

using the Hilbert-Schmidt inner product

$$\langle X,Y\rangle = \mathrm{Tr}(X^T Y).$$

If Y is a solution to (66), then integration by parts yields the remarkable equality

$$\mathrm{Tr}(Y)(r) = E(Y,R,A).$$

If we minimize E with respect to all three arguments Y, R, and A, then we both solve (66) and minimize the logarithmic derivative of γ. If we fix Y, then it is immediate from (67) and from the constraints that we should take $R = \kappa I$ and $A_k = 0$, *i.e.*, maximum curvature and flat mirrors.

8.3. **Multiple images.** Lemma 5.7 yields the following model.

LP Problem 8.1. *Given n, κ, A, V, and m, is there a symmetric, positive measure $\mu(\ell,\alpha,\beta)$ such that*

$$\alpha_*(\mu) = \int_{\ell,\beta} \mathrm{d}\mu = A\,\mathrm{d}z(\alpha)$$

$$\int_{\ell,\alpha,\beta} s_{n,\kappa}(\ell)\sec(\alpha)\sec(\beta)\,\mathrm{d}\mu \leqslant mA^2$$

$$\int_{\ell,\alpha,\beta} s_{n,\kappa}^{(-1)}(\ell)\big(\sec(\alpha)+\sec(\beta)\big)\,\mathrm{d}\mu \leqslant 2mAV$$

$$\int_{\ell,\alpha,\beta} s_{n,\kappa}^{(-2)}(\ell)\,\mathrm{d}\mu \leqslant mV^2$$

$$\int_{\ell,\alpha,\beta} \ell\,\mathrm{d}\mu = \omega_{n-1}V\,?$$

Theorem 1.8 now follows as a porism[5] of Theorem 1.4. If we apply the transformation

$$\tilde{V} = mV, \qquad \tilde{A} = mA, \qquad \tilde{\mu} = m\mu,$$

then Problem 8.1 becomes Problem 6.1.

8.4. **Alternative functionals.** In this section we prove Theorem 1.18. The proof is almost the same as the proof of Croke's theorems in Section 7.1.

Given $L = L(\Omega)$, we consider the following linear programming problem based on equation 25.

[5]A corollary of proof.

LP Problem 8.2. *Given n, A, and L, is there a symmetric positive measure* $\mu(\ell,\alpha,\beta)$ *such that*

$$\alpha_*(\mu) = \int_{\ell,\beta} \mathrm{d}\mu = A\,\mathrm{d}z(\alpha)$$

$$\int_{\ell,\alpha,\beta} \ell^{n-1}\sec(\alpha)\sec(\beta)\,\mathrm{d}\mu \leqslant A^2$$

$$\int_{\ell,\alpha,\beta} \ell^{n-3}\,\mathrm{d}\mu = L?$$

We can apply a version of Lemma 6.4 to establish Theorem 1.18 as a sharp inequality. Given $a \geqslant 0$ and $d \in \mathbb{R}$, we consider the cost function

$$E(\ell,\alpha,\beta) = a\ell^{n-1}\sec(\alpha)\sec(\beta) - d\ell^{n-3}.$$

By design, given a radius $r > 0$, there are values of $a, d > 0$ such that $E(\ell,\alpha,\alpha)$ is minimized in ℓ when

$$\ell = 2r\cos(\alpha),$$

which thus satisfies (52). We can take

$$a = \frac{n-3}{r^2}, \qquad d = 4(n-1).$$

(Note that we need $n \geqslant 4$. If $n < 3$, then a would be negative. If $n = 3$, then $L(\Omega) \propto |\partial\Omega|$ and Theorem 1.18 is vacuous.)

We define the potential

$$f(\alpha) = -\frac{E(2r\cos(\alpha),\alpha,\alpha)}{2} = 2^{n-1}(r\cos(\alpha))^{n-3}.$$

Applying the change of variables (54), the adjusted cost function is

$$F(\ell,x,y) = (n-3)\ell^{n-1}xy - 4(n-1)\ell^{n-3} + 2^{n-1}(x^{3-n} + y^{3-n}).$$

We want to show that $F \geqslant 0$. For any fixed value of xy, $F(\ell,x,y)$ is minimized when $x = y$. Then

$$\begin{aligned} F(\ell,x,x) &= (n-3)\ell^{n-1}x^2 - 4(n-1)\ell^{n-3} + 2^n x^{3-n} \\ &= \big((n-3)(\ell x)^{n-1} - 4(n-1)(\ell x)^{n-3} + 2^n\big)x^{3-n}. \end{aligned}$$

The first factor is a polynomial in ℓx that, by univariate calculus, decreases to 0 at $\ell x = 2$ and then increases again. This completes the proof of Theorem 1.18.

8.5. **Old wine in new decanters.** In this section, we complete the proof of Theorem 1.17. The rest of this paper has covered all cases except Theorem 1.16, due to Yau, and Theorem 1.15, due to Croke. The arguments given here are equivalent to the original proofs, only restated in linear programming form.

8.5.1. *Yau's linear isoperimetric inequality.* If Ω is n-dimensional and LCD(-1), then Problem 7.2 yields

$$\int_{\ell,\alpha,\beta} \big(s^{(-1)}(\ell)\sec(\alpha) - (n-1)s^{(-2)}(\ell)\big)\,\mathrm{d}\mu_\Omega \leqslant AV - (n-1)V^2$$

since $q > n-1$. The integrand is positive, since it is the second antiderivative of

$$s'(\ell)\sec(\alpha) - (n-1)s(\ell) = (n-1)\sinh(\ell)^{n-2}\big(\cosh(\ell)\sec(\alpha) - \sinh(\ell)\big) > 0.$$

Thus the right side is positive, and Theorem 1.16 follows. In terms of optimal transport, the result follows if we define a cost function

$$E(\ell,\alpha,\beta) = s^{(-1)}(\ell)\sec(\alpha) - (n-1)s^{(-2)}(\ell),$$

and then a vanishing potential $f(\alpha) = 0$.

8.5.2. *Croke's curvature-free inequality.* For simplicity, we take $\rho = 1$.

Suppose that Ω is n-dimensional with unique geodesics. Lemma 5.6 produces the following simple model independent of κ, and that can be combined with Problem 6.1.

LP Problem 8.3. *Given n, A, and V, is there a symmetric, positive measure $\mu(\ell,\alpha,\beta)$ such that*

$$\alpha_*(\mu) = \int_{\ell,\beta} \mathrm{d}\mu = A\,\mathrm{d}z(\alpha)$$
$$\int_{\ell,\alpha,\beta} s^{(-2)}_{n,(\pi/\ell)^2}(\ell)\,\mathrm{d}\mu \leqslant V^2$$
$$\int_{\ell,\alpha,\beta} \ell\,\mathrm{d}\mu = \omega_{n-1}V?$$

To analyze this model, we simplify it in two respects. First, we can integrate away α and β, because none of the integrals explicitly depend on them. We call the resulting measure $\mu(\ell)$. Second, we can explicitly evaluate the integrand that arises from Lemma 5.6:

$$s^{(-2)}_{n,(\pi/\ell)^2}(\ell) = \left(\frac{\ell}{\pi}\right)^{n+1} s^{(-2)}_{n,1}(\pi) = \frac{\ell^{n+1}\omega_n}{2\pi^n\omega_{n-1}}.$$

The first equality is just rescaling by ℓ/π. The second equality is a tricky but standard integral; the answer can also be inferred from the optimal case of a hemisphere $Y_{n,1}$. The simplified model is then as follows.

LP Problem 8.4. *Given n, A, and V, is there a positive measure $\mu(\ell)$ on $\mathbb{R}_{\geqslant 0}$ such that*

$$\int_\ell \mathrm{d}\mu = \frac{\omega_{n-2}}{n-1}A$$
$$\int_\ell \frac{\ell^{n+1}\omega_n}{2\pi^n\omega_{n-1}}\,\mathrm{d}\mu_\Omega \leqslant V^2$$
$$\int_\ell \ell\,\mathrm{d}\mu_\Omega = \omega_{n-1}V$$

As usual, we state the dual of Problem 8.4.

LP Problem 8.5. *Given n, A, and V, are there constants $c \geqslant 0$ and $f, d \in \mathbb{R}$ such that*

$$f + c\frac{\ell^{n+1}\omega_n}{2\pi^n\omega_{n-1}} - d\ell \geqslant 0$$
$$f\frac{\omega_{n-2}}{n-1}A + cV^2 - d\omega_{n-1}V < 0? \tag{68}$$

In the optimal case of $Y_{n,1}$, we have $\ell = \pi$ everywhere and $V = \omega_n/2$. We can solve for the constants f, c, and d assuming that the left side of equation (68) reaches 0 there and is non-negative for other values of ℓ. We obtain

$$c = 2\omega_{n-1}, \qquad d = (n+1)\omega_n, \qquad f = n\pi\omega_n.$$

Assuming that A is feasible for Problem 8.4, (68) then gives us the inequality

$$A \geqslant \frac{(n-1)\omega_n\omega_{n-1}}{2\pi\omega_{n-2}} = \omega_{n-1}.$$

This establishes Theorem 1.15.

Remark. It may seem wrong that $Y_{n,1}$ does not itself have unique geodesics. But it is a limit of manifolds that do, which is good enough. In any case the proof of Theorem 1.15 only really uses that Ω has unique geodesics in its interior.

9. Closing questions

Of course, we want Theorem 1.5 without the smallness condition (4). It would suffice to prove a stronger version of Lemma 5.5. This would be implied by the $n = 4$ case of the following conjecture.

Conjecture 9.1. *Let $j(r,t) = j_M(\gamma, r, t)$ be the candle function of a geodesic in γ in an n-manifold M with curvature $K \leqslant -1$. Then*

$$\Big[(n-1)^2 j - (n-1)\frac{\partial j}{\partial t} + (n-1)\frac{\partial j}{\partial r} - \frac{\partial^2 j}{\partial r \partial t}\Big](r,t) \tag{69}$$

is minimized when M has constant curvature $K = -1$.

As in the proof of Lemma 5.5, we would use Conjecture 9.1 to obtain the inequality

$$\int_{\ell,\alpha,\beta} \Big(\frac{s_{n,-1}(\ell)}{\cos(\alpha)\cos(\beta)} - (n-1)s^{(-1)}_{n,-1}(\ell)\big(\frac{1}{\cos(\alpha)} + \frac{1}{\cos(\beta)}\big) + (n-1)^2 s^{(-2)}_{n,-1}(\ell)\Big)\,\mathrm{d}\mu_\Omega$$
$$\leqslant |\partial\Omega|^2 - 2(n-1)|\partial\Omega||\Omega| + (n-1)^2|\Omega|^2,$$

and thus sharpen Problem 7.2, by integrating over $(\gamma\cap\Omega)\times(\gamma\cap\Omega)$ for a general geodesic γ. If we integrate over a connected interval $[0,\ell]$, which suffices when Ω is convex, then Conjecture 9.1 implies that

$$j(0,\ell) - (n-1)\int_0^\ell j(0,t)\,\mathrm{d}t - (n-1)\int_0^\ell j(s,\ell)\,\mathrm{d}s + (n-1)^2\int_0^\ell\int_0^t j(s,t)\,\mathrm{d}s\,\mathrm{d}t$$

is minimized when $K = -1$. Note that even this relation is not true under the weaker hypothesis $\mathrm{LCD}(-1)$. For example, it does not hold when ℓ is large enough if M is the complex hyperbolic plane $\mathbb{C}\mathrm{H}^2$, normalized to be $(-9/4, -9/16)$-pinched.

The following relaxation of Kleiner's theorem is open even though, as explained in Section 1.4.2, it is close to true. The motivation is that the even strongest form holds in dimension $n = 4$ following the proof of Croke's theorem.

Question 9.1. *Suppose that Ω is a compact 3-manifold with boundary, and with unique geodesics, non-positive curvature, and fixed volume $V = |\Omega|$. Then is its surface area $|\partial\Omega|$ minimized when Ω is a round, Euclidean ball? What if non-positive curvature is replaced by* $\mathrm{Candle}(0)$*? What if* $\mathrm{Candle}(0)$ *is only required for pairs of boundary points?*

Question 9.1 could also be asked in dimension $n \geqslant 5$ and for other curvature bounds $\kappa \neq 0$.

Finally, the following conjecture would give a more robust proof of Theorem 1.7, with a weaker hypothesis as well when $\kappa = 0$.

Conjecture 9.2. *Suppose that Ω is a convex, compact Riemannian n-manifold with boundary with unique geodesics. Suppose that for some constants κ and r, all chords in Ω satisfy equation (52). Then Ω is isometric to $B_{n,\kappa}(r)$.*

REFERENCES

[AB04] Stephanie B. Alexander and Richard L. Bishop, *Curvature bounds for warped products of metric spaces*, Geom. Funct. Anal. **14**:6 (2004), 1143–1181. 4.1, 4.1

[Ada05] Colin Adams, *Hyperbolic knots*, Handbook of knot theory, Elsevier, 2005, pp. 1–18, arXiv:math/0309466. 4.1

[Aub76] Thierry Aubin, *Problèmes isopérimétriques et espaces de Sobolev*, J. Differential Geometry **11**:4 (1976), 573–598. 1.1, 4

[BC64] Richard L. Bishop and Richard J. Crittenden, *Geometry of manifolds*, Pure and Applied Mathematics, vol. XV, Academic Press, 1964. 1.1, 2.2, 5.5, 8.1

[Ber03] Marcel Berger, *A panoramic view of Riemannian geometry*, Springer-Verlag, Berlin, 2003. 1.1

[BO69] Richard L. Bishop and Barrett O'Neill, *Manifolds of negative curvature*, Trans. Amer. Math. Soc. **145** (1969), 1–49. 4.1

[Bol41] Gerrit Bol, *Isoperimetrische Ungleichungen für Bereiche auf Flächen*, Jber. Deutsch. Math. Verein. **51** (1941), 219–257. 1.1, 1.11, 1.12

[BR33] Edwin F. Beckenbach and Tibor Radó, *Subharmonic functions and surfaces of negative curvature*, Trans. Amer. Math. Soc. **35**:3 (1933), 662–674. 1.1

[BZ88] Yuri Burago and Victor A. Zalgaller, *Geometric inequalities*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 285, Springer-Verlag, 1988, Translated from the Russian by A. B. Sosinskiĭ, Springer Series in Soviet Mathematics. 1.1

[CE03] Henry Cohn and Noam Elkies, *New upper bounds on sphere packings. I*, Ann. of Math. (2) **157**:2 (2003), 689–714, arXiv:math/0110009. 1.3

[Cha84] Isaac Chavel, *Eigenvalues in Riemannian geometry*, Pure and Applied Mathematics, vol. 115, Academic Press, 1984. 1.1

[Cho03] Jaigyoung Choe, *Relative isoperimetric inequality for domains outside a convex set*, Arch. Inequal. Appl. **1**:2 (2003), 241–250. 1.2.5, 1.14

[Cho06] Jaigyoung Choe, *The double cover relative to a convex domain and the relative isoperimetric inequality*, J. Aust. Math. Soc. **80**:3 (2006), 375–382. 1.2.5, 1.14, 5.5

[Cro80] Christopher B. Croke, *Some isoperimetric inequalities and eigenvalue estimates*, Ann. Sci. École Norm. Sup. (4) **13**:4 (1980), 419–435. 1.1, 1.15, 5.4

[Cro84] Christopher B. Croke, *A sharp four-dimensional isoperimetric inequality*, Comment. Math. Helv. **59**:2 (1984), 187–192. 1.1, 1.1, 1.13, 1.4.2, 5, 5.3

[CS93] John H. Conway and Neil J. A. Sloane, *Sphere packings, lattices and groups*, 3rd ed., Grundlehren der mathematischen Wissenschaften, vol. 290, Springer-Verlag, New York, 1993. 1.3

[Del72] Philippe Delsarte, *Bounds for unrestricted codes, by linear programming*, Philips Res. Rep. **27** (1972), 272–289. 1.3

[Dru02] Olivier Druet, *Sharp local isoperimetric inequalities involving the scalar curvature*, Proc. Amer. Math. Soc. **130**:8 (2002), 2351–2361 (electronic). 1.1

[Dru10] Olivier Druet, *Isoperimetric inequalities on nonpositively curved spaces*, 2010, Notes from intensive course at Queen Dido Conference, May 2010, Carthage, Tunisia, `http://math.arizona.edu/~dido/presentations/Druet-Carthage.pdf`. 1.1, 4

[Far01] Julius Farkas, *Theorie der einfachen Ungleichungen*, J. Reine Angew. Math. **124** (1901), 1–27. 6.2.1

[GHL90] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine, *Riemannian geometry*, second ed., Universitext, Springer-Verlag, 1990. 8.2

[Gro81] Mikhael Gromov, *Structures métriques pour les variétés riemanniennes*, Textes Mathématiques, vol. 1, CEDIC, Paris, 1981, Edited by J. Lafontaine and P. Pansu. 1.1, 1.1

[Gro99] Misha Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics, vol. 152, Birkhäuser Boston Inc., 1999, Based on the 1981 French original. 1.1

[Gün60] Paul Günther, *Einige Sätze über das Volumenelement eines Riemannschen Raumes*, Publ. Math. Debrecen **7** (1960), 78–93. 1.1, 2.2, 5.5

[Has94] Joel Hass, *Bounded* 3*-manifolds admit negatively curved metrics with concave boundary*, J. Differential Geom. **40**:3 (1994), 449–459. 1.2.6, 4

[Has16] Joel Hass, *Isoperimetric regions in nonpositively curved manifolds*, 2016, arXiv:1604.02768. 4

[HHM99] Hugh Howards, Michael Hutchings, and Frank Morgan, *The isoperimetric problem on surfaces*, Amer. Math. Monthly **106**:5 (1999), 430–439. 3.1

[Inn98] Nobuhiro Innami, *Integral formulas for polyhedral and spherical billiards*, J. Math. Soc. Japan **50**:2 (1998), 339–357. 8.2

[Kje02] Tinne Hoff Kjeldsen, *Different motivations and goals in the historical development of the theory of systems of linear inequalities*, Arch. Hist. Exact Sci. **56**:6 (2002), 469–538. 6.2.1

[KK15] Benoît Kloeckner and Greg Kuperberg, *A refinement of Günther's candle inequality*, Asian J. Math **19**:1 (2015), 121–134, arXiv:1204.3943. 1.1, 2.2, 2.2

[Kle92] Bruce Kleiner, *An isoperimetric comparison theorem*, Invent. Math. **108**:1 (1992), 37–47. 1.1, 1.2.6

[Klo15] Benoît R. Kloeckner, *Curvatures and anisometry of maps*, Comm. Anal. Geom. **23**:2 (2015), 319–348, arXiv:1403.4197. 1.1

[McD03] Kirk T. McDonald, *Maximal gravity at the surface of an asteroid*, 2003, arXiv:physics/0312029. 3.1

[Min10] Hermann Minkowski, *Geometrie der Zahlen*, B. G. Teubner Verlag, 1910. 6.2.1

[MJ00] Frank Morgan and David L. Johnson, *Some sharp isoperimetric theorems for Riemannian manifolds*, Indiana Univ. Math. J. **49**:3 (2000), 1017–1041. 1.1, 4

[MS98] Dusa McDuff and Dietmar Salamon, *Introduction to symplectic topology*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, 1998. 5

[OS79] A. M. Odlyzko and N. J. A. Sloane, *New bounds on the number of unit spheres that can touch a unit sphere in n dimensions*, J. Combin. Theory Ser. A **26**:2 (1979), 210–214. 1.3

[Oss78] Robert Osserman, *The isoperimetric inequality*, Bull. Amer. Math. Soc. **84**:6 (1978), 1182–1238. 1.1

[PV16] Stefano Pigola and Giona Veronelli, *The smooth riemannian extension problem*, 2016, arXiv:1606.08320. 1.2.6

[Rit05] Manuel Ritoré, *Optimal isoperimetric inequalities for three-dimensional Cartan-Hadamard manifolds*, Global theory of minimal surfaces, Clay Math. Proc., vol. 2, Amer. Math. Soc., 2005, pp. 395–404. 1.1

[Sage] *Sage open-source mathematical software system*, `http://www.sagemath.org/`. 1.3

[San04] Luis A. Santaló, *Integral geometry and geometric probability*, second ed., Cambridge Mathematical Library, Cambridge University Press, 2004. 5, 5.3

[Smi07] Warren Smith, *Modern optical engineering*, fourth ed., McGraw-Hill Professional, 2007. 5.2

[Teu91] Eberhard Teufel, *A generalization of the isoperimetric inequality in the hyperbolic plane*, Arch. Math. (Basel) **57**:5 (1991), 508–513. 5.3

[Teu93] Eberhard Teufel, *On integral geometry in Riemannian spaces*, Abh. Math. Sem. Univ. Hamburg **63** (1993), 17–27. 5

[Vil09] Cédric Villani, *Optimal transport, old and new*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, 2009. 6.2.2

[Wei26] André Weil, *Sur les surfaces a courbure negative*, C. R. Acad. Sci. Paris **182** (1926), 1069–1071. 1.1, 1.2.3, 1.11

[Yau75] Shing-Tung Yau, *Isoperimetric constants and the first eigenvalue of a compact Riemannian manifold*, Ann. Sci. École Norm. Sup. (4) **8**:4 (1975), 487–507. 1.16, 5.2

E-mail address: `benoit.kloeckner@u-pec.fr`

UNIVERSITÉ PARIS-EST, LABORATOIRE D'ANALYSE ET DE MATÉMATIQUES APPLIQUÉES (UMR 8050), UPEM, UPEC, CNRS, F-94010, CRÉTEIL, FRANCE

E-mail address: `greg@math.ucdavis.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, DAVIS