Rev. Mat. Iberoam. 38 (2022), no. 3, 705–729
DOI 10.4171/RMI/1309

Rigidity of the Pu inequality and quadratic isoperimetric constants of normed spaces

Paul Creutz

Abstract. Our main result gives an improved bound on the filling areas of curves in Banach spaces which are not closed geodesics. As applications we show rigidity of Pu's classical systolic inequality and investigate the isoperimetric constants of normed spaces. The latter has further applications concerning the regularity of minimal surfaces in Finsler manifolds.

1. Introduction

1.1. Rigidity of the Pu inequality

Let $d\colon S^1 \times S^1 \to \mathbb{R}$ be a metric on the circle. Then the *filling area* of d is defined as

$$\mathrm{Fill}(d) := \inf_g \{\mathrm{Area}(g)\},$$

where g ranges over all Riemannian metrics g on the disc D^2 such that the boundary distance function $\mathrm{bd}_g\colon S^1 \times S^1 \to \mathbb{R}$ satisfies $\mathrm{bd}_g \geq d$. We call a Riemannian metric g on D^2 a *minimal filling* if its area equals the filling area of its boundary distance function. These definitions can be generalized to higher dimensions and more general surfaces. In any case it is usually difficult to find criteria which imply that a given metric is a minimal filling. Compare [5,9,12,13,23,27,30] for some positive results of this type. One concrete example of a minimal filling is the round metric on the hemisphere H^2. Hence

$$\mathrm{Fill}(d_{S^1}) = 2\pi,$$

where d_{S^1} denotes the angular metric on S^1. As has been noted in [23], the latter is equivalent to Pu's sharp systolic inequality for the projective plane, see [43]. One of our main results is the rigidity of the Pu inequality in the following sense.

Theorem 1.1. *Let d be a metric on S^1 such that $d \leq d_{S^1}$ and $d \neq d_{S^1}$. Then*

$$\mathrm{Fill}(d) < 2\pi.$$

2020 Mathematics Subject Classification: Primary 53C24; Secondary 46B20, 28A75, 49Q05, 30L05.
Keywords: Isoperimetric inequality, systolic inequality, normed spaces, minimal surfaces, area functionals.

Theorem 1.1 might seem non surprising at first glance. However, it contrasts the non-rigidity of the Besicovitch inequality, [8], which has been noted in [14]. The reason for this non-rigidity is that not all metrics on S^1 arise as boundary distance functions of Riemannian metrics on the disc. Indeed it seems a challenging problem to describe the ones that do arise this way. Only in the more flexible Finsler setting some understanding has been provided in [14].

More generally we study filling areas of curves in a Banach space X. Here a subtle issue becomes important. Namely that of defining the area of two dimensional subsets of X, or equivalently of X-valued maps from the disc. One possibility is to consider the standard Hausdorff 2-measure $\mathcal{H}_X^2$ and define the *Busemann area functional* $\mathcal{A}^b$ by setting

$$\mathcal{A}^b(f) := \int_X \operatorname{card}\left(f^{-1}(y)\right) \mathrm{d}\mathcal{H}_X^2(y). \tag{1.1}$$

for Lipschitz discs $f\colon D^2 \to X$. However, depending on the context, other definitions sometimes appear to be more natural, see e.g. [2]. The most popular ones are the Benson–Gromov mass* area functional $\mathcal{A}^{m*}$, commonly used in geometric measure theory due to its strong convexity properties, and the Holmes–Thompson area functional $\mathcal{A}^{\mathrm{ht}}$, which has certain desirable properties from the point of view of Finsler geometry. There is a natural list of common properties that all relevant examples share, and one calls an object satisfying all of them an *area functional* (in the sense of convex geometry), see Section 2.2 below. All area functionals agree up to a universal multiplicative constant, and thus in many contexts the concrete choice of area functional does not play a major role. However, for the quantitative type of questions that we are studying in this article, it is important at several points. Thus in the following we will fix an abstract area functional $\mathcal{A}$ and specify further when it becomes necessary.

The *filling area* $\operatorname{Fill}^{\mathcal{A}}(\gamma)$ of a closed curve γ in X is defined as the infimum of $\mathcal{A}(f)$, where f ranges over all Lipschitz discs spanning γ. The main result of [16] states: *if $\gamma\colon (S^1, d_{S^1}) \to X$ is 1-Lipschitz, then γ extends to a 1-Lipschitz map $G\colon H^2 \to X$.* We refine this result as follows.

Theorem 1.2. *If γ is 1-Lipschitz but not an isometric embedding, then G is area decreasing in the sense that $\mathcal{A}(G) < 2\pi$. In particular, $\operatorname{Fill}^{\mathcal{A}}(\gamma) < 2\pi$.*

It follows from [28] that Theorem 1.1 is in fact a special case of Theorem 1.2. Roughly speaking, it corresponds to choosing $X = \ell^\infty$ and $\mathcal{A} = \mathcal{A}^{\mathrm{ir}}$ as Ivanov's inscribed Riemannian area functional. Some other applications of Theorem 1.2 are discussed in Sections 1.2 and 1.3 below. Beyond these, Theorem 1.2 seems also amenable to future applications in cut and paste arguments in the context of systolic inequalities and similar problems.

1.2. Quadratic isoperimetric spectra

The isoperimetric profile, or (geometric) *Dehn function* $\delta_X^{\mathcal{A}}\colon (0,\infty) \to [0,\infty]$, of a metric space X is defined by

$$\delta_X^{\mathcal{A}}(r) := \sup_\gamma \left\{ \operatorname{Fill}^{\mathcal{A}}(\gamma) \right\},$$

where γ ranges over all closed Lipschitz curves in X such that $\ell(\gamma) \le r$. If X is simply connected and satisfies some weak geometric assumptions, then the asymptotic growth of the geometric Dehn function is the same as that of the combinatorial Dehn function of a group acting geometrically on X, see [24, 40]. The latter is a well-studied quasi-isometry invariant in geometric group theory. The *isoperimetric spectrum* IP is defined as the set of those α in $[1,\infty)$ such that there is is a finitely presentable group with asymptotic growth $\simeq r^\alpha$. By [24] and [10], its closure is given as

$$\overline{\mathrm{IP}} = \{1\} \cup [2,\infty). \tag{1.2}$$

The gap in (1.2) extends into the case of quadratic growth by the following result of Wenger, [51]: *if X is a proper geodesic metric space such that*

$$\limsup_{r\to\infty} \frac{\delta_X^b(r)}{r^2} < \frac{1}{4\pi}, \tag{1.3}$$

then X is Gromov hyperbolic, and hence the asymptotic of its Dehn function is in fact even linear. Note that here we denote $\delta_X^b := \delta_X^{\mathcal{A}^b}$ and that in the following we will use this type of notational simplifications without further mentioning. This result is sharp as the Dehn function of the Euclidean space $\mathbb{R}^n$ is given by $\frac{1}{4\pi}\cdot r^2$, independently of $\mathcal{A}$, see Example 4.4. The implications of a non-strict inequality in (1.3) have been investigated in [52].

In the present paper we study the following finer non-coarse quantity:

$$C^{\mathcal{A}}(X) := \sup_{r\in(0,\infty)} \frac{\delta_X^{\mathcal{A}}(r)}{r^2} \in [0,\infty].$$

We call $C^{\mathcal{A}}(X)$ the *$\mathcal{A}$-quadratic isoperimetric constant* of X. Its investigation may be motivated by the following remarkable result due to Lytchak–Wenger, [39]: *A proper geodesic metric space X is* CAT(0) *if and only if*

$$C^b(X) \in \left\{0, \frac{1}{4\pi}\right\}.$$

Compare also the proof of Theorem 5.4 in [40] for the gap between 0 and $\frac{1}{4\pi}$.

The *$\mathcal{A}$-quadratic isoperimetric spectrum* $\mathrm{QIS}^{\mathcal{A}}(\mathbf{M})$ of a class of metric spaces $\mathbf{M}$ is the set of $\mathcal{A}$-quadratic isoperimetric constants of its elements. Improving Theorem 1.2 in [16], we are able to give a full description of the quadratic isoperimetric spectrum of the class of all Banach spaces.

Theorem 1.3. *Let* **Ban** *be the class of Banach spaces. Then*

$$\mathrm{QIS}^b(\mathbf{Ban}) = \{0\} \cup \left[\frac{1}{4\pi}, \frac{1}{2\pi}\right]. \tag{1.4}$$

For **Ban**, the quadratic isoperimetric spectrum is the same for the most reasonable choices of area functional $\mathcal{A}$, and is given by (1.4). However, for the class $\mathbf{Ban}_n$ of normed spaces of fixed finite dimension n, the quadratic isoperimetric spectrum very much depends on the choice of area functional. For $n = 2$, the spectra of the aforementioned functionals may be determined from classical results in convex geometry as:

	$\mathcal{A} = \mathcal{A}^{\mathrm{ht}}$	$\mathcal{A} = \mathcal{A}^b$	$\mathcal{A} = \mathcal{A}^{m*}$	$\mathcal{A} = \mathcal{A}^{\mathrm{ir}}$
$\mathrm{QIS}^{\mathcal{A}}(\mathbf{Ban}_2) =$	$\left\{\frac{1}{4\pi}\right\}$	$\left[\frac{1}{4\pi}, \frac{\pi}{32}\right]$	$\left[\frac{1}{4\pi}, \frac{1}{8}\right]$	$\left[\frac{1}{4\pi}, \frac{1}{8}\right]$

See Example 4.8. In particular, every two dimensional normed space satisfies the Euclidean isoperimetric inequality with respect to $\mathcal{A}^{\mathrm{ht}}$, while it satisfies the Euclidean isoperimetric inequality with respect to $\mathcal{A}^{b}$ only if it is Euclidean. The following result clarifies the behaviour between dimensions 2 and ∞.

Theorem 1.4. *Let $n \geq 2$ and let $\mathcal{A}$ be an area functional such that $\mathcal{A} \geq \mathcal{A}^{\mathrm{ht}}$. Then the $\mathcal{A}$-quadratic isoperimetric spectrum of* $\mathbf{Ban}_n$ *is a compact interval $[\frac{1}{4\pi}, r_n^{\mathcal{A}}]$, where $r_n^{\mathcal{A}} < \frac{1}{2\pi}$ is nondecreasing in n and converges to $\frac{1}{2\pi}$ as $n \to \infty$.*

The assumption $\mathcal{A} \geq \mathcal{A}^{\mathrm{ht}}$ is satisfied for $\mathcal{A} = \mathcal{A}^{b}, \mathcal{A}^{\mathrm{ht}}, \mathcal{A}^{m*}, \mathcal{A}^{\mathrm{ir}}$, see Section 2.2. Explicit values of the optimal constants $r_n^{\mathcal{A}}$ beyond the aforementioned case $n = 2$ remain open. It is natural to think of our setting as the isoperimetric problem in dimension one. In the case $n = 2$ we benefit of the coincidence of dimension one and *co*dimension one. The latter is the mostly studied situation, and it is essentially solved in finite dimensional normed spaces as well as many other classes of spaces, see for example [3, 21, 32, 41, 45]. Beyond dimension one and codimension one, isoperimetric inequalities have been obtained in [4, 23, 50]. However, sharp constants are only known in the Euclidean space and very few other situations, see [1, 46].

1.3. Minimal surfaces in Finsler manifolds

Let X be a proper metric space which satisfies a local quadratic isoperimetric inequality and let Γ be a rectifiable Jordan curve in X. Set $\Lambda(\Gamma, X)$ to be the set of those Sobolev discs $u \in W^{1,2}(D^2, X)$ for which the trace $u_{|S^1}$ gives a monotone parametrization of Γ. The following solution of the Plateau problem has been given by Lytchak–Wenger in [36, 37]: *if $\Lambda(\Gamma, X) \neq \emptyset$, then there is $u \in \Lambda(\Gamma, X)$ of least parametrized Hausdorff measure which moreover may be chosen infinitesimally isotropic.* Such u will be called a *solution of the Plateau problem*. Variants of the metric space valued Plateau problem have been solved for collections of Jordan curves, surfaces of higher genus and self-intersecting curves in [17, 18, 22].

For a solution u of Plateau's problem, a factorization $u = \bar{u} \circ P$ with the following properties has been investigated in [38]:

- Z_u is a geodesic metric space homeomorphic to D^2,
- $P\colon D^2 \to Z_u$ is monotone,
- and $\bar{u}\colon D^2 \to X$ is 1-Lipschitz.

An analytically more well-behaved variation of this factorization has been discussed in [20]. A *branch point* of u is a point in D^2 where u is not a local embedding. In general the set of branch points may be large and the map P may be non-injective, see Example 11.3 in [38]. However, Question 11.4 in [38] asks: *Can the set of branch points of a solution of the Plateau problem be large if the isoperimetric constant C is smaller than $\frac{1}{2\pi}$? Can the map P be non-injective in this case?* By Theorem 1.4, a positive answer to this question would apply in the case that X is a finite dimensional normed space or a compact Finsler manifold. This would be desirable as up to now the set of branch points of solutions of the Plateau problem in Finsler manifolds can only be controlled under restrictive assumptions on X and Γ, see Theorem 1.6 in [42].

Remark 1.5. Since this article first appeared as a preprint, I have shown in [19] together with Matthew Romney that no such control on the set of branch points is possible when only assuming the strict bound on the quadratic isoperimetric constant. However, the counterexample ambient spaces constructed in [19] are in other senses geometrically quite degenerate. Hence there is still a reasonable chance that the results of this article can be an ingredient in a proof showing a control on the set of branch points in Finsler manifolds.

The quadratic isoperimetric constant of X also controls the Hölder regularity of solutions of the Plateau problem and more general X-valued (quasi-)harmonic discs, see [35,36]. In particular, Theorem 1.4 may be applied to improve the α-Hölder regularity of solutions of the Plateau problem in Finsler manifolds calculated in Theorem 1.4 of [16] beyond the threshold case to $\alpha > \pi/8$. Similar calculations lead to concrete uniform Hölder constants for minimal surfaces in Finsler manifolds in the settings of [35], [22], [17] and [18].

1.4. Outline of proof and byproducts

In this subsection we shortly discuss the main ideas entering in the proofs of Theorems 1.2 and 1.4. We start with the proof of Theorem 1.2. For sake of simplicity we restrict here to the case that X is finite dimensional. Let $p \in H^2$ be a point of differentiability of G and let $v \in T_p H^2$ and $q, \bar{q} \in S^1$ be the endpoints of the great arc passing through p in direction v. If $\|d_p G(v)\| = |v|$, then there is $\Lambda \in X^*$ satisfying $\|\Lambda\| = 1$ and $\Lambda(d_p G(v)) = |v|$. A somewhat analytical argument involving the optimal transport going on in the proof of the main result of [16] then shows that

$$(\Lambda \circ \gamma)(\bar{q}) - (\Lambda \circ \gamma)(q) = \pi,$$

and hence γ restricted to $\{q, \bar{q}\}$ is isometric. By assumption this cannot hold for all q, and in particular G must be infinitesimally shrinking at p in some direction. If $\mathcal{A}$ is the Busemann or the Holmes–Thompson area functional, then this implies the claim since both these area functionals are strictly monotone. The proof for general Banach spaces and area functionals that we perform below is conceptually similar but more technical.

To prove Theorem 1.4 we endow $\mathbf{Ban}_n$ with the Banach–Mazur distance. Then $C^{\mathcal{A}}(X)$ is continuous in X and hence the quadratic isoperimetric spectrum of $\mathbf{Ban}_n$ is a compact interval $[l_n^{\mathcal{A}}, r_n^{\mathcal{A}}]$. It then follows from [26] and [11] that $l_n^{\mathrm{ht}} = l_2^{\mathrm{ht}} = \frac{1}{4\pi}$. Hence if $\mathcal{A} \geq \mathcal{A}^{\mathrm{ht}}$ then $l_n^{\mathcal{A}} = \frac{1}{4\pi}$. In order to show

$$r_n^{\mathcal{A}} < \frac{1}{2\pi}, \tag{1.5}$$

we prove the existence of extremal curves in the following sense.

Lemma 1.6. *Let X be a finite dimensional normed space of dimension ≥ 2. Then there is a bi-Lipschitz embedding $\gamma\colon S^1 \to X$ satisfying*

$$\mathrm{Fill}^{\mathcal{A}}(\gamma) = C^{\mathcal{A}}(X) \cdot \ell(\gamma)^2.$$

Note that at least for $\mathcal{A} = \mathcal{A}^{\mathrm{ht}}$ counterintuitively these extremal curves can be planar only if $C^{\mathrm{ht}}(X) = \frac{1}{4\pi}$. The concrete shape of such curves remains mysterious except for very particular cases, see Example 4.4. To prove (1.5), we choose $X \in \mathbf{Ban}_n$ such that $C^{\mathcal{A}}(X) = r_n^{\mathcal{A}}$ and within X an $\mathcal{A}$-extremal curve γ. Without loss of generality we may assume that $\ell(\gamma) = 2\pi$ and γ is 1-Lipschitz. Inequality (1.5) is then implied by Theorem 1.2 and the following lemma.

Lemma 1.7. *Let X be a finite dimensional normed space. Then there is no isometric embedding of (S^1, d_{S^1}) into X.*

The proof of Lemma 1.7 relies on an explicit description of geodesics in X in terms of the structure of its unit ball that we give below. From this characterization it follows that if γ is an isometric embedding, then the derivative $\gamma' \colon S^1 \to X$ would be a measurable function which is 'too' discontinuous.

Non surprisingly, it is a hard task to give lower bounds on the filling areas of curves. Our main tool at hand is a generalization of the Pu inequality due to Sergei Ivanov, which implies: *if $\gamma \colon (S^1, d_{S^1}) \to X$ is an isometric embedding, then*

$$\operatorname{Fill}^{\mathrm{ht}}(\gamma) \geq 2\pi, \tag{1.6}$$

see [28,29]. Lemma 1.7 seemingly indicates that (1.6) cannot be applied. However, we can still embed isometrically large finite portions of S^1 into $\mathbb{R}^n_\infty$. Such embeddings together with a homotopy argument invoking (1.6) imply

$$r_n^{\mathcal{A}} \geq r_n^{\mathrm{ht}} \geq C^{\mathrm{ht}}(\mathbb{R}^n_\infty) \geq \left(1 - \frac{4}{n}\right) \frac{1}{2\pi}. \tag{1.7}$$

Note that the lower bound (1.7) leading to the asymptotic behaviour of $r_n^{\mathcal{A}}$ is explicit, while the upper bound (1.5) is obtained by contradiction.

1.5. Organization

In Section 2 we recall some basic facts and set up notation. First, in Section 2.1, we state a characterization of the John ellipse that will be needed in the proof of Theorem 1.2. Then, in Section 2.2 we recall the notion of area functionals and discuss different examples and comparison results between them. Finally, in Section 2.3 we discuss some basic homotopy arguments and their applications. Section 3 is dedicated to the proofs of Theorem 1.1 and 1.2. In Section 3.1 we recall the construction of the extension map G from [16] and to this end also the relevant optimal transport theory. In the subsequent Section 3.2 we show a somewhat technical lemma on the dependence of certain optimal transport plans on the base point. This lemma will be needed in Section 3.3, where we proof a crucial proposition that allows us to relate global properties of the curve γ to the infinitesimal behaviour of the corresponding extension map G. Finally, in Section 3.4 we are then able to give the proofs of Theorems 1.1 and 1.2. In the remaining Section 4 we perform the proof of Theorem 1.4. First, in Sections 4.1 and 4.2, we prove Lemma 1.7 and 1.6 respectively. Then in Section 4.3 we discuss the quadratic isoperimetric spectra of $\mathbf{Ban}_n$ for general area functionals. Finally, we complete the proofs of Theorems 1.3 and 1.4 in Section 4.4 by discussing lower bounds such as (1.7).

2. Preliminaries

2.1. The John ellipse

Let B be a compact, convex, centrally symmetric subset of $\mathbb{R}^2$ which contains the origin in its interior. Then the John ellipsoid theorem states that there is a unique ellipse E of maximal volume contained in B. This ellipse is called the *John ellipse* and satisfies $E \subseteq B \subseteq \sqrt{2} \cdot E$. We will need the following characterization.

Lemma 2.1. *Let $(X, |\cdot|)$ be an Euclidean plane and let $\|\cdot\|$ be a norm on X which satisfies $\|\cdot\| \leq |\cdot|$. Denote by E the unit ball of $|\cdot|$ and by B the unit ball of $\|\cdot\|$. Then E is the John ellipse of B if and only if there exist $v_0, v_1, v_2 \in X$ such that $|v_i| = \|v_i\| = 1$ and*

$$\langle v_i, v_{i+1} \rangle \geq 0, \quad i = 0, 1, 2, \tag{2.1}$$

where we define $v_3 = \tau(v_0) := -v_0$ as the antipodal point of v_0.

Note that Lemma 2.1 implies that if E is the John ellipse of B, then there are vectors $v, w \in X$ with

$$|v| = |w| = \|v\| = \|v\| = 1$$

such that the angle between v and w lies in the interval $[\pi/6, \pi/2]$.

Proof of Lemma 2.1. We may assume without loss of generality that $X = \mathbb{R}^2$ and that $|\cdot|$ is the standard Euclidean norm. Then D^2 is the John ellipse of B if and only if there are $v_0, v_1, v_2 \in \mathbb{R}^2$ such that $|v_i| = \|v_i\| = 1$ and $\lambda_0, \lambda_1, \lambda_2 \geq 0$ for which

$$I_2 = \lambda_0 \cdot v_0 \otimes v_0 + \lambda_1 \cdot v_1 \otimes v_1 + \lambda_2 \cdot v_2 \otimes v_2, \tag{2.2}$$

see [25]. Here $I_2 \in \mathrm{GL}_2$ denotes the identity matrix. We show that the latter condition is equivalent to (2.1). In either case we may assume $v_0 = (1, 0)$ and furthermore that v_0, v_1, v_2 are cyclically ordered, contained in the upper half plane and pairwise distinct. Set $a_0 := d_{S^1}(v_1, v_2)$, $a_1 := d_{S^1}(v_2, \tau(v_0))$ and $a_2 := d_{S^1}(v_0, v_1)$. Then (2.2) becomes

$$\begin{aligned} I_2 = \lambda_0 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} &+ \lambda_1 \cdot \begin{pmatrix} \cos^2(a_2) & \frac{1}{2}\sin(2a_2) \\ \frac{1}{2}\sin(2a_2) & \sin^2(a_2) \end{pmatrix} \\ &+ \lambda_2 \cdot \begin{pmatrix} \cos^2(a_1) & -\frac{1}{2}\sin(2a_1) \\ -\frac{1}{2}\sin(2a_1) & \sin^2(a_1) \end{pmatrix}. \end{aligned}$$

Note that $a_0, a_2 \in (0, \pi)$ and $a_1 \in [0, \pi)$. Thus if $a_1 \notin \{0, \pi/2\}$ then solving this system of equations gives

$$\lambda_i = \frac{\sin(2a_i)}{\sin(2a_1)\sin^2(a_2) + \sin(2a_2)\sin^2(a_1)}$$

for $i = 0, 1, 2$. Note here that our assumptions on the vectors v_0, v_1, v_2 imply that at most one of the terms $\sin(2a_i)$ can be nonpositive. Hence in the case $a_1 \notin \{0, \pi/2\}$ all the λ_i are nonnegative if and only if (2.1) holds.

If $a_1 = \pi/2$ and hence $v_2 = (0, 1)$, then certainly condition (2.1) is satisfied, and the same is true for (2.2) upon choosing $\lambda_0 = \lambda_2 = 1$ and $\lambda_1 = 0$. Finally, assume $a_1 = 0$ and hence $v_2 = (-1, 0)$. Then both conditions (2.1) and (2.2) respectively imply $v_1 = (0, 1)$ and hence that the respective other one is satisfied. ∎

2.2. Area functionals

The aim of this subsection is to shortly discuss area functionals in the sense of convex geometry. We will follow the approach of [37] based on Jacobians. The reader is referred to [3, 7, 28] for other equivalent viewpoints.

Let Σ be the set of seminorms on $\mathbb{R}^2$ and let Σ_0 be the set of norms on $\mathbb{R}^2$.

Definition 2.2. A *Jacobian* is a map $\mathbf{J}\colon \Sigma \to [0, \infty)$ fulfilling the following properties:

(1) (Normalization) $\mathbf{J}(|\cdot|) = 1$ for $|\cdot|$ the standard Euclidean norm on $\mathbb{R}^2$.

(2) (Monotonicity) $\mathbf{J}(s) \geq \mathbf{J}(s')$ whenever $s \geq s'$.

(3) (Transformation law) $\mathbf{J}(s \circ T) = |\det T|\, \mathbf{J}(s)$ for $T \in M_2(\mathbb{R})$.

Example 2.3. It follows readily that $\mathbf{J}(s) = 0$ if and only if s is degenerate. So it suffices to define the following examples of Jacobians on a norm $\|\cdot\| \in \Sigma_0$ with unit ball B.

(1) The *Busemann Jacobian* $\mathbf{J}^{\mathrm{b}}$ is defined by

$$\mathbf{J}^{\mathrm{b}}\big(\|\cdot\|\big) := \frac{\pi}{\mathcal{L}^2(B)},$$

where $\mathcal{L}^2$ denotes the standard Lebesgue measure on $\mathbb{R}^2$.

(2) The *Holmes–Thompson Jacobian* $\mathbf{J}^{\mathrm{ht}}$ is defined by

$$\mathbf{J}^{\mathrm{ht}}\big(\|\cdot\|\big) := \frac{\mathcal{L}^2(B^*)}{\pi},$$

where $B^* := \{v \in \mathbb{R}^2 |\ \langle v, w\rangle \leq 1; \forall w \in B\}$ is the polar body of B.

(3) The *Benson–Gromov mass* Jacobian* is defined by

$$\mathbf{J}^{m*}\big(\|\cdot\|\big) := \sup_P \frac{4}{\mathcal{L}^2(P)},$$

where P ranges over all parallelograms containing B.

(4) *Ivanov's inscribed Riemannian Jacobian* $\mathbf{J}^{\mathrm{ir}}$ is defined by

$$\mathbf{J}^{\mathrm{ir}}\big(\|\cdot\|\big) := \frac{\pi}{\mathcal{L}^2(E)}, \tag{2.3}$$

where E is the John ellipse of B. Dually, the *circumscribed Riemannian Jacobian* $\mathbf{J}^{\mathrm{cr}}$ is defined by taking E in (2.3) as a ellipse of least area containing B.

Remark 2.4. The presented Jacobians satisfy the following comparison results.

(1) From the Blaschke–Santaló inequality and the definitions we deduce that

$$\mathbf{J}^{\mathrm{cr}}\big(\|\cdot\|\big) \leq \mathbf{J}^{\mathrm{ht}}\big(\|\cdot\|\big) \leq \mathbf{J}^{\mathrm{b}}\big(\|\cdot\|\big) \leq \mathbf{J}^{\mathrm{ir}}\big(\|\cdot\|\big), \tag{2.4}$$

where each of the inequalities is strict if and only if $\|\cdot\|$ is not Euclidean.

(2) Dually one can see that

$$\mathbf{J}^{\mathrm{ir}}\big(\|\cdot\|\big) \leq \frac{4}{\pi}\mathbf{J}^{\mathrm{b}}\big(\|\cdot\|\big) \leq \frac{\pi}{2}\mathbf{J}^{\mathrm{ht}}\big(\|\cdot\|\big) \leq 2 \cdot \mathbf{J}^{\mathrm{cr}}\big(\|\cdot\|\big), \tag{2.5}$$

where equality is attained if and only if B is a parallelogram. The middle part is the Mahler–Reisner inequality, see [48]. The other two follow from [37] and duality.

(3) More generally, $\mathbf{J}^{\mathrm{ir}}$ is maximal among all Jacobians, and $\mathbf{J}^{\mathrm{cr}}$ is minimal. For a Jacobian $\mathbf{J}$ we define

$$q^{\mathbf{J}} := \inf_{\|\cdot\| \in \Sigma_0} \frac{\mathbf{J}(\|\cdot\|)}{\mathbf{J}^{\mathrm{ir}}(\|\cdot\|)} \in \Big[\frac{1}{2}, 1\Big].$$

By (2.5) we have $q^{\mathrm{ir}} = 1$, $q^{\mathrm{ht}} = 2/\pi$, $q^{\mathrm{b}} = \pi/4$ and $q^{\mathrm{cr}} = 1/2$. From Lemma 2.1 one may deduce that $q^{m*} = \sqrt{3}/2$ is attained if B is a regular hexagon.

(4) For the Holmes–Thompson and the mass* area functionals we have

$$\frac{2}{\pi}\mathbf{J}^{m*}(\|\cdot\|) \le \mathbf{J}^{\mathrm{ht}}(\|\cdot\|) \le \mathbf{J}^{m*}(\|\cdot\|), \tag{2.6}$$

where equality on the left is attained if and only if B is a parallelogram, and equality on the right is attained if and only if $\|\cdot\|$ is Euclidean. This follows from [3] and the previous observations.

In the following, let $U \subset \mathbb{R}^2$ be open, let X be a metric space, and let $f : U \to X$ be a Lipschitz map. At almost every $p \in U$, the *metric differential* $\mathrm{md}_p f$ is well defined as a seminorm on $\mathbb{R}^2$ via

$$(\mathrm{md}_p f)(v) := \lim_{t \to 0} \frac{d(f(p + tv), f(p))}{|t|}.$$

Indeed the metric differential transforms nicely with respect to coordinate changes and thus one can define metric differentials also in the case that S is a smooth 2-dimensional manifold and $g : S \to X$ is a Lipschitz map, compare e.g. [22], p. 93. For a Jacobian $\mathbf{J}$, we define the corresponding *area functional* $\mathcal{A}^{\mathbf{J}}$ by setting

$$\mathcal{A}^{\mathbf{J}}(f) := \int_U \mathbf{J}(\mathrm{md}_p f)\, \mathrm{d}\mathcal{L}^2(p). \tag{2.7}$$

For $\mathcal{A}^b$, equation (2.7) is consistent with equation (1.1) by a variant of the area formula, see [31].

The definition of Jacobians is cooked up as to obtain the following natural list of properties for the arising area functionals.

Lemma 2.5. (1) (*Normalization*) *If X is a Riemannian manifold, then $\mathcal{A}(f) = \mathcal{A}^b(f)$.*

(2) (*Monotonicity*) *If $g : X \to Y$ is L-Lipschitz, then $\mathcal{A}(g \circ f) \le L^2 \cdot \mathcal{A}(f)$.*

(3) (*Coordinate invariance*) *If $V \subseteq \mathbb{R}^2$ is a open and $\varphi : V \to U$ is bi-Lipschitz, then $\mathcal{A}(f \circ \varphi) = \mathcal{A}(f)$.*

Remark 2.6. By the coordinate invariance, one can naturally extend area functionals to assign areas to Lipschitz maps $f : M \to X$, where M is a smooth 2-dimensional manifold.

2.3. Homotopy arguments

In this section we fix an area functional $\mathcal{A}$ and a metric space X.

For a Lipschitz curve $\gamma: S^1 \to X$, we define its $\mathcal{A}$-*filling area* by

$$\operatorname{Fill}^{\mathcal{A}}(\gamma) := \inf\left\{\mathcal{A}(f) \mid f: \bar{D}^2 \to X \text{ Lipschitz}, f_{|S^1} = \gamma\right\}.$$

The crucial observation for homotopy arguments is: *if $h: S^1 \times [0,1] \to X$ is a Lipschitz map, then*

$$|\operatorname{Fill}^{\mathcal{A}}(\gamma_0) - \operatorname{Fill}^{\mathcal{A}}(\gamma_1)| \leq \mathcal{A}(h),$$

where $\gamma_i = h(\cdot, i)$. The following lemma allows to restrict to curves which are parametrized by constant-speed in most situations.

Lemma 2.7. *Let $\gamma_0, \gamma_1: S^1 \to X$ be Lipschitz curves and reparametrizations of each other. Then*

$$\operatorname{Fill}^{\mathcal{A}}(\gamma_1) = \operatorname{Fill}^{\mathcal{A}}(\gamma_2).$$

Proof. By Lemma 3.6 in [40], there exists a Lipschitz homotopy h between γ_0 and γ_1 such that $\mathcal{A}(h) = 0$. ■

The following simple but useful lemma will be applied various times.

Lemma 2.8. *Let X be a geodesic metric space, $\gamma_0, \gamma_1: S^1 \to X$ closed Lipschitz curves and $\phi_0, \dots, \phi_m \in S^1$ cyclically ordered points. Then*

$$|\operatorname{Fill}^{\mathcal{A}}(\gamma_0) - \operatorname{Fill}^{\mathcal{A}}(\gamma_1)| \leq C \sum_{k=0}^{m} (l_k^0 + l_k^1 + d_k + d_{k+1})^2,$$

where $C := C^{\mathcal{A}}(X)$, $d_k := d(\gamma_0(\phi_k), \gamma_1(\phi_k))$ and $l_k^i := \ell(\gamma_{i\,|[\phi_k,\phi_{k+1}]})$.

Proof. We define the Lipschitz homotopy $h: S^1 \times [0,1] \to X$ between γ_0 and γ_1 by setting $h(\phi_k, \cdot)$ to be a geodesic connecting $\gamma_0(\phi_k)$ to $\gamma_1(\phi_k)$ and filling the remaining squares by application of the quadratic isoperimetric inequality. ■

For $L \geq 0$, let $\Gamma^L(X)$ be the set of closed L-Lipschitz curves in X endowed with the maximum metric

$$d_\infty(\gamma_0, \gamma_1) := \max_{\phi \in S^1} d(\gamma_0(\phi), \gamma_1(\phi)).$$

Corollary 2.9. *If X is geodesic and $C := C^{\mathcal{A}}(X) < \infty$, then $\operatorname{Fill}^{\mathcal{A}}: \Gamma^L(X) \to \mathbb{R}$ is continuous.*

Proof. Let $\varepsilon := d_\infty(\gamma_0, \gamma_1) \leq \pi L$. Choosing $m := \lceil \pi L/\varepsilon \rceil$ equidistant points on S^1 and applying Lemma 2.8 gives

$$|\operatorname{Fill}^{\mathcal{A}}(\gamma_0) - \operatorname{Fill}^{\mathcal{A}}(\gamma_1)| \leq Cm\left(\frac{4\pi L}{m} + 2\varepsilon\right)^2 \leq 72\pi CL \cdot \varepsilon.$$

Compare also the proof of Lemma 18 in [47]. ■

3. Proofs of Theorems 1.1 and 1.2

3.1. The extension map

The following extension theorem has been obtained in [16].

Theorem 3.1. *Let X be a Banach space and let $\gamma\colon S^1 \to X$ be 1-Lipschitz. Then γ extends to a 1-Lipschitz map $G\colon H^2 \to X$.*

Remember here that H^2 denotes the upper hemisphere endowed with the round metric. Because we will work with the explicit definition of the extension map G, we now recall its construction, which relies on optimal transport theory.

Let (X, d) be a complete metric space. Denote by $\mathcal{P}(X)$ the set of separably supported Borel probability measures on X and by $\mathcal{P}_1(X)$ the set of those $\mu \in \mathcal{P}(X)$ satisfying

$$\int_X d(x, y)\, \mathrm{d}\mu(y) < \infty$$

for some $x \in X$. For a continuous map $f\colon X \to Y$, denote by $f_*\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ the push forward map given by $f_*\mu(A) = \mu(f^{-1}(A))$. For $\mu, \nu \in \mathcal{P}(X)$, we call $K \in \mathcal{P}(X \times X)$ a *coupling* from μ to ν if $\pi_{1*}K = \mu$ and $\pi_{2*}K = \nu$. Denote the set of couplings from μ to ν by $\Pi(\mu, \nu)$. The metric space obtained by endowing $\mathcal{P}_1(X)$ with the distance

$$d_W(\mu, \nu) := \inf_{K \in \Pi(\mu,\nu)} \int_{X \times X} d(x, y)\, \mathrm{d}K(x, y)$$

will be called the *Wasserstein-1-space over X* and will also shortly be denoted by $\mathcal{P}_1(X)$. A measurable map $T\colon X \to X$ will be called an *optimal transport plan* from μ to ν if $T_*\mu = \nu$ and

$$d_W(\mu, \nu) = \int_X d(x, T(x))\, \mathrm{d}\mu(x).$$

The proof of Theorem 3.1 relies on four things. First, for every complete metric space there is a canonical isometric embedding δ of X into $\mathcal{P}_1(X)$ given by mapping $x \in X$ to the Dirac measure δ_x. Secondly, if X is a Banach space, then there is a 1-Lipschitz retraction $b\colon \mathcal{P}_1(X) \to X$ given by

$$b(\mu) := \int_X x\, \mathrm{d}\mu(x).$$

The third is that if $f\colon X \to Y$ is 1-Lipschitz then also the push forward map $f_*\colon \mathcal{P}_1(X) \to \mathcal{P}_1(Y)$ is 1-Lipschitz. The last but most important observation is the following.

Proposition 3.2. *There is an isometric embedding $\mu\colon H^2 \to \mathcal{P}_1(S^1)$ extending the Dirac embedding $\delta\colon S^1 \to \mathcal{P}_1(S^1)$.*

Theorem 3.1 follows from these observations by setting $G\colon H^2 \to X$ to be given by

$$G := b \circ \gamma_* \circ \mu.$$

To construct the embedding μ, it is important to understand optimal transport plans in dimension one. For intervals we have the following simple description, cf. [49].

Lemma 3.3. *Let $I \subset \mathbb{R}$ be a closed interval and let $\mu, \nu \in \mathcal{P}_1(I)$ be absolutely continuous measures with strictly increasing distribution functions $F_\mu, F_\nu \colon I \to [0,1]$. Then an optimal transport map from μ to ν is given by $T = F_\nu^{-1} \circ F_\mu$ and*

$$d_W(\mu,\nu) = \int_I |F_\mu(s) - F_\nu(s)|\,\mathrm{d}s = \int_0^1 |F_\mu^{-1}(t) - F_\nu^{-1}(t)|\,\mathrm{d}t. \tag{3.1}$$

We are however mainly interested in understanding the Wasserstein distance on S^1. This problem has been solved by Cabrelli–Molter in [15], and we shortly discuss their approach here.

For a point $\phi \in S^1$ and absolutely continuous $\mu \in \mathcal{P}^1(S^1)$, we denote by μ^ϕ the measure on $[0, 2\pi]$ which corresponds to μ under the orientation preserving 'identification' of $[0, 2\pi]$ and S^1 mapping 0 to ϕ. Then for $\mu, \nu \in \mathcal{P}_1(S^1)$, the inequality

$$d_W(\mu,\nu) \le d_W(\mu^\phi, \nu^\phi)$$

is immediate. For $\mu, \nu \in \mathcal{P}_1(S^1)$, we call $\phi \in S^1$ an *equilibrated cutpoint* for (μ,ν) if there is a Borel partition $[0, 2\pi) = A \dot\cup B$ such that $|A| = |B|$, $F_{\mu^\phi} \le F_{\nu^\phi}$ on A and $F_{\mu^\phi} \ge F_{\nu^\phi}$ on B. This definition is justified by the following theorem.

Theorem 3.4 ([15]). *Let $\mu, \nu \in \mathcal{P}_1(S^1)$ be absolutely continuous measures. Then there exists an equilibrated cutpoint ϕ for (μ,ν), and for every such ϕ one has*

$$d_W(\mu,\nu) = d_W(\mu^\phi, \nu^\phi).$$

So calculating the Wasserstein distance between distributions on S^1 amounts to finding an equilibrated cutpoint and then calculating the integral (3.1).

The construction of the map μ in Proposition 3.2 goes as follows. For fixed $p \in H^2 \setminus S^1$, let

$$d_p \colon S^1 \to \mathbb{R}, \psi \mapsto d_{S^2}(p, \psi)$$

be the distance function to p. Let $b_p \in S^1$ be such that $d_{S^2}(p, b_p) = d_{S^2}(p, S^1)$ and

$$k_p := \cos(d_{S^2}(p, b_p)).$$

Let $h_p \colon S^1 \to \mathbb{R}$ be given by

$$h_p(\psi) := \frac{1}{2}(d_p''(\psi))^+ + \frac{1-k_p}{2\pi},$$

and let μ_p be the measure on S^1 which is absolutely continuous with density h_p. The most technical part in the proof that $\mu \colon H^2 \to \mathcal{P}_1(S^1)$ defines an isometric embedding amounts to the following lemma, compare Sections 3.2 and 3.3 in [16].

Lemma 3.5. *Let $\eta \colon [0,\pi] \to H^2$ be a great arc not contained in S^1 and set $\phi := \eta(0) \in S^1$.*

- *If $r, s \in (0, \pi)$, then ϕ is a balanced cutpoint for $(\mu_{\eta(s)}, \mu_{\eta(r)})$.*
- *If furthermore $r \le s \le \pi/2$, then a corresponding Borel partition is given by $A = [0,\pi)$ and $B = [\pi, 2\pi]$.*

Note that in the statement of Lemma 3.5 we implicitly assume that the great arc η is parametrized by arc-length and that in the following we will always make this implicit assumption on great arcs and great circles.

3.2. Variation of optimal transport plans

For the proof of Theorem 1.2 we will need the following technical observation concerning the dependence of the optimal transport plan from μ_p to μ_q on the points $p, q \in H^2$.

Lemma 3.6. *Let $\eta: [0,\pi] \to H^2$ be a great arc not contained in S^1 and let $0 < r < \pi/2$. Set $p := \eta(r)$, $v := \eta'(r)$ and $\phi := \eta(0)$. Then there is a C^1-map $T = T(s,\psi): (0,\pi/2) \times S^1 \to S^1$ such that, for every fixed s,*

(1) *the map $T(s,\cdot)$ is an optimal transport plan from $\mu_{\eta(r)}$ to $\mu_{\eta(s)}$,*

(2) *$\frac{\partial T}{\partial s}(s,\cdot)$ is positive almost everywhere on the interval $[\phi, \tau(\phi)]$ and negative almost everywhere on $[\tau(\phi), \phi]$.*

Here as usual we denote by τ the antipodal map of S^1.

Proof. Set $\mu_s := \mu_{\eta(s)}$, $d_s := d_{\eta(s)}$, $k_s := k_{\eta(s)}$, $h_s := h_{\eta(s)}$, $b_s := b_{\eta(s)}$.

Identify S^1 and $[0, 2\pi]$ such that ϕ corresponds to 0. We set $\mathcal{D} := (0,\pi/2) \times [0, 2\pi]$ and define the analytic function $d: \mathcal{D} \to \mathbb{R}$ by $d(s,\psi) := d_s(\psi)$. Furthermore, we define $F: \mathcal{D} \to [0,1]$ by

$$F(s,\psi) = F_s(\psi) := \mu_s([0,\psi)) = \int_0^{\psi} h_s(\varphi)\,\mathrm{d}\varphi.$$

Then by Lemma 3.3, Theorem 3.4 and Lemma 3.5, an optimal transport map T^s from μ_r to μ_s is given by

$$T^s(\psi) = F_s^{-1}(F_r(\psi)).$$

Furthermore, if $s \geq t$, then $T^s(\psi) \geq T^t(\psi)$ for $\psi \in [0,\pi]$ and $T^s(\psi) \leq T^t(\psi)$ for $\psi \in [\pi, 2\pi]$. We define the map $T: \mathcal{D} \to [0, 2\pi]$ by $T(s,\psi) := T^s(\psi)$.

Let ξ be the angle between the great arc η and the circle S^1 in the point ϕ. Then by the spherical sine theorem we have that

$$k(s) := k_s = \sqrt{1 - \sin^2(s)\sin^2(\xi)}$$

is an analytic function with nonzero derivative on $(0,\pi/2)$. We have $b_s \in (-\pi/2, \pi/2)$ and hence by Lemma 3.1 in [16], for $\psi \in [0, 2\pi]$ that $d_s''(\psi) \leq 0$ when

$$b_s + \frac{\pi}{2} \leq \psi \leq b_s + \frac{3\pi}{2},$$

and $d_s''(\psi) \geq 0$ otherwise. Thus by the fundamental theorem of calculus and the first variation formula,

$$F(s,\psi) = \frac{1}{2}\cdot \begin{cases} \frac{\partial d}{\partial \psi}(s,\psi) - \cos(\xi) + \psi\,\frac{1-k_s}{\pi}, & 0 \leq \psi \leq b_s + \frac{\pi}{2}, \\ \frac{\partial d}{\partial \psi}\big(s, b_s + \frac{\pi}{2}\big) - \cos(\xi) + \psi\,\frac{1-k_s}{\pi}, & b_s + \frac{\pi}{2} \leq \psi \leq b_s + \frac{3\pi}{2}, \\ \frac{\partial d}{\partial \psi}(s,\psi) + 2\,\frac{\partial d}{\partial \psi}\big(s, b_s + \frac{\pi}{2}\big) - \cos(\xi) + \psi\,\frac{1-k_s}{\pi}, & b_s + \frac{3\pi}{2} \leq \psi \leq 2\pi. \end{cases}$$

By Lemma 3.1 in [16],

$$\frac{\partial^2 d}{(\partial \psi)^2}\Big(s, b_s + \frac{\pi}{2}\Big) = 0,$$

implying that $\partial F/\partial s$ is well defined and continuous on $\mathcal{D}$. Also

$$\frac{\partial F}{\partial \psi}(s,\psi) = h_s(\psi)$$

is continuous on $\mathcal{D}$ and hence $F \in C^1(\mathcal{D})$. By the first variation formula,

$$\frac{\partial d}{\partial \psi}\Big(s, b_s + \frac{\pi}{2}\Big) = k_s.$$

So for fixed parameter s, the function $\frac{\partial F}{\partial s}(s,\cdot)\colon [0,2\pi] \to \mathbb{R}$ is piecewise analytic on the intervals $\big[0, b_s+\pi/2\big], [b_s+\pi/2, b_s+3\pi/2], [b_s+3\pi/2, 2\pi]$, and on $[b_s+\pi/2, b_s+3\pi/2]$ it is given by

$$\frac{\partial F}{\partial s}(s,\psi) = \frac{1}{2}\cdot\frac{\partial k}{\partial s}(s)\cdot\Big(1-\frac{\psi}{\pi}\Big),$$

which is zero only for $s=\pi$. In particular $\frac{\partial F}{\partial s}(s,\cdot)$ has only finitely many zeros.

For fixed s, the map $F_s\colon [0,2\pi] \to [0,1]$ defines a C^1-diffeomorphism and hence by the implicit function theorem,

$$\frac{\partial}{\partial s}\big(F_s^{-1}(v)\big) = \frac{-1}{\frac{\partial F}{\partial \psi}(s, F_s^{-1}(v))}\cdot\frac{\partial F}{\partial s}\big(s, F_s^{-1}(v)\big).$$

Hence T is differentiable in s and

$$\frac{\partial T}{\partial s}(s,\psi) = -\frac{\frac{\partial F}{\partial s}(s, T^s(\psi))}{h_s(T^s(\psi))}$$

is continuous on $\mathcal{D}$. In particular $\frac{\partial T}{\partial s}(s,\psi)=0$ if and only if $\frac{\partial F}{\partial s}(s,T^s(\psi))=0$, implying that $\frac{\partial T}{\partial s}(s,\cdot)$ has only finitely many zeros. To complete the proof that T is C^1, it suffices to note that

$$\frac{\partial T}{\partial \psi}(s,\psi) = \frac{-h_r(\psi)}{h_s(T^s(\psi))}$$

is continuous on $\mathcal{D}$. ∎

3.3. Local versus global

A key observation for the proof of Theorem 1.2 is the following proposition. It will allow us to conclude that if the curve γ is 1-Lipschitz but not isometric then also, at almost every point, the extension map G must be infinitesimally shrinking in certain directions.

Proposition 3.7. *Let $\eta\colon [0,\pi] \to H^2$ be a great arc not contained in S^1 and let $r \neq 0, \pi/2, \pi$. Set $p := \eta(r)$, $v := \eta'(r)$ and $\phi := \eta(0)$. If p is a point of metric differentiability of G and $(\mathrm{md}_p\, G)(v) = 1$, then*

$$\|\gamma(\phi) - \gamma(\tau(\phi))\| = \pi,$$

where $\tau\colon S^1 \to S^1$ is the antipodal map.

Proof. Set $\mu_s := \mu_{\eta(s)}$ and let T be as in Lemma 3.6. Then

$$(G\circ\eta)(s) = \int_{S^1} \gamma(\psi)\,\mathrm{d}\mu_s(\psi) = \int_{S^1} \gamma(T(s,\psi))\,\mathrm{d}\mu_r(\psi). \tag{3.2}$$

By assumption there is a sequence $s_n \searrow r$ such that

$$1-\frac{1}{n} \le \frac{\|(G\circ\eta)(s_n)-(G\circ\eta)(r)\|}{s_n-r}. \tag{3.3}$$

Choose $\Lambda_n \in X^*$ such that $\|\Lambda_n\| = 1$ and

$$\Lambda_n\big((G\circ\eta)(s_n)-(G\circ\eta)(r)\big) = \|(G\circ\eta)(s_n)-(G\circ\eta)(r)\|. \tag{3.4}$$

Then by (3.2), (3.3), (3.4) and the fundamental theorem of calculus for Lipschitz functions, we obtain

$$\begin{aligned} 1-\frac{1}{n} &\le \frac{1}{s_n-r}\Lambda_n\Big(\int_{S^1}\gamma(T(s_n,\psi))-\gamma(\psi)\,\mathrm{d}\mu_r(\psi)\Big)\\ &= \frac{1}{s_n-r}\int_{S^1}\int_r^{s_n}(\Lambda_n\circ\gamma)'(T(s,\psi))\cdot\frac{\partial T}{\partial s}(s,\psi)\,\mathrm{d}s\,\mathrm{d}\mu_r(\psi). \end{aligned}$$

In particular there is $t_n \in [r, s_n]$ such that

$$1-\frac{1}{n} \le \int_{S^1} \underbrace{(\Lambda_n\circ\gamma)'\big(T(t_n,\psi)\big)}_{=:f_n(\psi)}\cdot\underbrace{\frac{\partial T}{\partial s}(t_n,\psi)}_{=:g_n(\psi)}\,\mathrm{d}\mu_r(\psi). \tag{3.5}$$

Then as $|f_n| \le 1$ and g_n converges uniformly,

$$\int_{S^1} f_n\cdot g_n\,\mathrm{d}\mu_r \le \int_{S^1}|g_n|\,\mathrm{d}\mu_r \overset{n\to\infty}{\longrightarrow} \int_{S^1}\underbrace{\Big|\frac{\partial T}{\partial s}(r,\psi)\Big|}_{=:g(\psi)}\,\mathrm{d}\mu_r(\psi). \tag{3.6}$$

As μ is an isometric embedding and $T(s,\cdot)$ an optimal transport plan,

$$\int_{S^1} g(\psi)\,\mathrm{d}\mu_r(\psi) = \int_{S^1}\lim_{s\searrow r}\frac{|T(s,\psi)-\psi|}{s-r}\,\mathrm{d}\mu_r(\psi) = \lim_{s\searrow r}\frac{d_W(\mu_s,\mu_r)}{s-r} = 1. \tag{3.7}$$

By (3.5), (3.6) and (3.7), $f_n\cdot g_n \to g$ in μ_r-measure. So up to passing to a subsequence, $f_n\cdot g_n \to g$ holds μ_r-almost everywhere and hence also $\mathcal{H}^1$-almost everywhere. Hence by Lemma 3.6, $f_n \to 1$ almost everywhere on $[\phi,\tau(\phi)]$. As the diffeomorphisms $T(t_n,\cdot)$ and their inverses are of uniformly bounded C^1 norm, $(\Lambda_n\circ\gamma)' \to 1$ in $\mathcal{H}^1$-measure on $[\phi,\tau(\phi)]$ and hence, up to again passing to a further subsequence, the convergence holds almost everywhere on $[\phi,\tau(\phi)]$. So by the fundamental theorem of calculus and the dominated convergence theorem,

$$\|\gamma(\phi)-\gamma(\tau(\phi))\| \ge \Lambda_n\big(\gamma(\tau(\phi))-\gamma(\phi)\big) = \int_\phi^{\tau(\phi)}(\Lambda_n\circ\gamma)'(\psi)\,\mathrm{d}\psi \overset{n\to\infty}{\longrightarrow} \pi,$$

which completes the proof. ∎

3.4. Proofs of Theorems 1.1 and 1.2

We now restate and then proof Theorem 1.2.

Theorem 1.2. *Let X be a Banach space and let $\mathcal{A}$ be an area functional. Furthermore let $\gamma\colon (S^1, d_{S^1}) \to X$ be a 1-Lipschitz curve and let $G\colon H^2 \to X$ be its 1-Lipschitz extension. If γ is not an isometric embedding, then $\mathcal{A}(G) < 2\pi$ and hence* $\mathrm{Fill}^{\mathcal{A}}(\gamma) < 2\pi$.

Proof. So assume that γ is not an isometric embedding. By its maximality it suffices to prove Theorem 1.2 for the inscribed Riemannian area functional $\mathcal{A}^{\mathrm{ir}}$. For $p \in H^2$, let $|\cdot|_p$ be the standard norm on $T_p H^2$ and let E_p be its unit ball. If p is a point of metric differentiability of G, then we denote by $B_p \subset T_p H^2$ the unit ball of the seminorm $\mathrm{md}_p G$. As $\mathrm{md}_p G \leq |\cdot|_p$, it suffices to prove that the set of points $p \in H^2$ for which E_p is not the John ellipse of B_p has positive measure.

As γ is not an isometric embedding, there exist $\phi_0 \in S^1$ and an open interval $I \subset S^1$ containing ϕ_0 such that, for all $\phi \in I$,

$$\|\gamma(\phi) - \gamma(\tau(\phi))\| < \pi, \tag{3.8}$$

where again $\tau\colon S^1 \to S^1$ denotes the antipodal map. We may assume that $\ell(I) < \pi/4$ and will denote $A := S^1 \setminus (I \cup \tau(I))$. For $p \in H^2$ and $\phi \in S^1$, let $\eta_{\phi,p}$ be a great arc passing through p which starts at ϕ and let $v(\phi, p) \in T_p H^2$ be the direction of $\eta_{\phi,p}$ in p. If $p \in H^2 \setminus S^1$, then $v(\cdot, p)$ defines a diffeomorphism between S^1 and the unit vectors in the tangent space at p.

Define $\alpha\colon A \times A \times \left(H^2 \setminus A\right) \to [0, \pi]$ by setting $\alpha(\phi, \psi, p)$ to be the angle between $v(\phi, p)$ and $v(\psi, p)$. Then α defines a continuous function and

$$\alpha(A \times A \times \{\phi_0\}) \subset \{0, \pi\}.$$

Thus the continuity of α and the compactness of A imply that there is $0 < \delta < \pi/4$ such that

$$\alpha(\phi, \psi, p) \notin \left[\frac{\pi}{6}, \frac{\pi}{2}\right]$$

whenever $\phi, \psi \in A$ and $p \in H^2$ satisfies $d_{S^2}(p, \phi_0) < \delta$.

We claim that E_p is not the John ellipse of B_p at every point $p \in H^2 \setminus S^1$ of metric differentiability of G such that $d_{S^2}(p, \phi_0) < \delta$. Since these points form a set of positive measure, this implies Theorem 1.2. To prove the claim, we let $p \in H^2 \setminus S^1$ be a point of metric differentiability such that $d_{S^2}(p, \phi_0) < \delta$. If $\phi \in S^1$ is such that

$$(\mathrm{md}_p G)(v(\phi, p)) = 1,$$

then by Proposition 3.7 either

$$\|\gamma(\phi) - \gamma(\tau(\phi))\| = \pi \quad \text{or} \quad d_{S^1}(\phi, \phi_0) \in \left[\frac{\pi}{2} - \delta, \frac{\pi}{2} + \delta\right].$$

By (3.8) and our assumptions that $\ell(I) < \pi/4$ and $\delta < \pi/4$, in both cases we may conclude that $\phi \in A$. Thus our choice of δ implies that if $w_1, w_2 \in T_p H^2$ are unit vectors satisfying

$$\mathrm{md}_p G(w_1) = \mathrm{md}_p G(w_2) = 1,$$

then the angle between w_1 and w_2 does not lie in the interval $[\pi/6, \pi/2]$. Hence the remark subsequent to the statement of Lemma 2.1 allows us to conclude that E_p is not the John ellipse of B_p. ■

Now, as promised in the introduction, Theorem 1.1 follows as a corollary.

Proof of Theorem 1.1. Let d be a metric on S^1 such that $d \leq d_{S^1}$ and $d \neq d_{S^1}$. We need to show that $\mathrm{Fill}(d) < 2\pi$. Denote by $X = L^\infty(\mathcal{H}^1_{S^1})$ the Banach space of $\mathcal{H}^1_{S^1}$-measurable functions $f: S^1 \to \mathbb{R}$ such that

$$\|f\|_\infty := \operatorname{esssup} |f| < \infty.$$

Then there is an isometric embedding $\iota: (S^1, d) \to X$ which is given by

$$(\iota(\phi))(\psi) = d(\phi, \psi),$$

see also Section 4.4 below. By our assumption on d, we have that the curve ι is 1-Lipschitz with respect to d_{S^1}, but not an isometric embedding. Hence Theorem 1.2 implies $\mathrm{Fill}^{\mathrm{ir}}(\iota) < 2\pi$. This completes the proof since by Corollary 5.7.1 in [28] one has

$$\mathrm{Fill}(d) = \mathrm{Fill}^{\mathrm{ir}}(\iota). \tag{3.9}$$

Note that in the statement of Corollary 5.7.1 in [28], both when filling metrics and when filling curves, the infima are taken over arbitrary compact surfaces which bound S^1 and not only over disc type ones. However, the proof therein goes by showing the equality of the infima separately for each fixed topological type of filling, see the proof of Theorem 5.6 right above the statement of Corollary 5.7 in [28]. In particular, the proof of Corollary 5.7.1 in [28] indeed shows (3.9). ■

4. Quadratic isoperimetric spectra

4.1. Geodesics in finite dimensional normed spaces

In this subsection we fix a finite dimensional normed space $(X, \|\cdot\|)$. The following lemma characterizes geodesics in X in terms of the shape of the unit ball.

Lemma 4.1. *Let $\gamma: [a, b] \to X$ be a 1-Lipschitz curve connecting the points p and q. Set $v := q - p$ and let $\Lambda \in X^*$ be such that $\|\Lambda\| = 1$ and $\Lambda(v) = \|v\|$. Then γ is an isometric embedding if and only if $\Lambda(\gamma'(t)) = 1$ for almost every $t \in [a, b]$.*

Proof. When applying the fundamental theorem of calculus to the 1-Lipschitz function $\Lambda \circ \gamma: [a, b] \to \mathbb{R}$, we get

$$\|v\| = \Lambda(v) = (\Lambda \circ \gamma)(b) - (\Lambda \circ \gamma)(a) = \int_a^b \Lambda(\gamma'(t))\,\mathrm{d}t \leq b - a.$$

So γ is an isometric embedding if and only if $\|v\| = b - a$, if and only if $\Lambda(\gamma'(t)) = 1$ almost everywhere in $[a, b]$. ■

The next aim is to prove Lemma 1.7 from the introduction.

Lemma 1.7. *There is no isometric embedding of (S^1, d_{S^1}) into X.*

For the proof we remind the reader that a metric space valued function $f: \mathbb{R}^m \to Y$ is called *approximately continuous* at $x \in \mathbb{R}^m$ if, for every $\varepsilon > 0$,

$$\lim_{r \downarrow 0} \frac{\mathcal{L}^m\big(\{y \in B_r(x) : d(f(y), f(x)) \geq \varepsilon\}\big)}{\mathcal{L}^m(B_r(x))} = 0.$$

If Y is a separable metric space, then a Borel measurable function $f: \mathbb{R}^m \to Y$ is approximately continuous almost everywhere, see [34].

Proof of Lemma 1.7. Let $\phi, \psi \in S^1$ be antipodal points such that $\|\gamma(\phi) - \gamma(\psi)\| = \pi$. Let v and Λ be chosen as in Lemma 4.1 for $p := \gamma(\phi)$ and $q := \gamma(\psi)$. Then it follows that γ is the composition of a shortest path γ_1 connecting p to q and a shortest path γ_2 connecting q to p. By Lemma 4.1, we have $\Lambda(\gamma_1'(t)) = 1$ almost everywhere and $\Lambda(\gamma_2'(t)) = -1$ almost everywhere. In particular, the measurable function $\gamma': S^1 \to X$ cannot be approximately continuous neither at ϕ nor at ψ. As X is separable, this implies the claim. ∎

Remark 4.2. Clearly Lemma 1.7 fails for general Banach spaces. Beyond the Kuratowski embedding of S^1 into ℓ^∞, it is also easy to write down an isometric embedding of S^1 into L^1. Note however that the proof of Lemma 1.7 goes through as soon as X is a Banach space which has the Radon–Nikodym property such as ℓ^1, cf. [6], Chapter 5.

4.2. Extremal curves

In this subsection we fix an area functional $\mathcal{A}$ and a finite dimensional normed space $(X, \|\cdot\|)$ of dimension at least two.

Recall that a closed Jordan curve γ in X is said to satisfy a *chord-arc condition* with constant $\lambda \geq 1$ if for every distinct $v, w \in \operatorname{im}(\gamma)$, the shorter of the two arcs of γ between v and w has length bounded above by $\lambda \cdot \|v - w\|$. A Jordan curve is bi-Lipschitz to S^1 if and only if it satisfies a chord-arc condition with *some* constant $\lambda \geq 1$. Next we prove the following quantitative version of Lemma 1.6.

Lemma 4.3. *There is a non-constant curve $\gamma: S^1 \to X$ such that*

$$\operatorname{Fill}^{\mathcal{A}}(\gamma) = C^{\mathcal{A}}(X) \cdot \ell(\gamma)^2.$$

Every such curve γ is a Jordan curve and satisfies a chord-arc condition with the constant $\lambda = \sqrt{2} + 1$.

Proof. First we note that $C := C^{\mathcal{A}}(X) \in (0, \infty)$. Indeed Theorem 1.2 in [16] implies $C^{\mathcal{A}}(X) \leq \frac{1}{2\pi}$. Furthermore, Theorems 4.4.1 and 4.4.2 in [48] together with the remark after the statement of Theorem 3 in [11] give that $C^{\mathrm{ht}}(X) \geq \frac{1}{4\pi}$. Hence (2.5) and the minimality of $\mathcal{A}^{\mathrm{cr}}$ imply

$$C \geq C^{\mathrm{cr}}(X) \geq \frac{\pi}{4} C^{\mathrm{ht}}(X) \geq \frac{1}{16}.$$

For a closed and nonconstant Lipschitz curve γ in X we define

$$I(\gamma) := \frac{\operatorname{Fill}^{\mathcal{A}}(\gamma)}{\ell(\gamma)^2}.$$

By definition there exists a sequence of closed nonconstant Lipschitz curves γ_n such that $I(\gamma_n) \nearrow C$. Note that $I(\gamma)$ remains unchanged when translating γ by a vector $v \in X$ or rescaling γ by a positive real $c \in (0, \infty)$. Thus we may assume that $\ell(\gamma_n) = 1$ and that the image of the γ_n is contained in the unit ball B for all n. Furthermore, thanks to Lemma 2.7, we know that the filling area of a Lipschitz curve does not depend on its parametrization and therefore we may assume that the curves γ_n are parametrized by constant speed and hence that they are 1-Lipschitz. We may even suppose that the sequence γ_n converges uniformly to a closed Lipschitz curve γ thanks to the Arzelà–Ascoli theorem. A normed vector space is geodesic and we have observed that $C < \infty$. Therefore Corollary 2.9 grants us with the continuity of the filling area map, which implies

$$I(\gamma_n) = \mathrm{Fill}^{\mathcal{A}}(\gamma_n) \to C = \mathrm{Fill}^{\mathcal{A}}(\gamma).$$

By lower semi-continuity of length, $\ell(\gamma) \leq 1$ and hence $I(\gamma) = C$.

Let γ be such that $I(\gamma) = C$ and assume γ is not Jordan or does not satisfy a chord-arc condition with constant λ. Then there exist $\phi_1, \phi_2 \in S^1$ such that the following holds: *if $l_1 \leq l_2$ are the lengths of the two arcs of γ connecting $\gamma(\phi_1)$ to $\gamma(\phi_2)$, then $\lambda \cdot d < l_1$, where $d := \|\gamma(\phi_1) - \gamma(\phi_2)\|$.*

Then by $l_1 \leq l_2$ and the particular choice of λ we have

$$(l_1 + d)^2 + (l_2 + d)^2 < l_1^2 + l_2^2 + \Big(\frac{4}{\lambda} + \frac{2}{\lambda^2}\Big) l_1 l_2 \leq (l_1 + l_2)^2 = \ell(\gamma)^2.$$

Applying Lemma 2.8, where we take $\gamma_0 := \gamma$ and γ_1 as the curve identically constant $\gamma(\phi_1)$ gives

$$\mathrm{Fill}^{\mathcal{A}}(\gamma) \leq C\,((l_1 + d)^2 + (l_2 + d)^2) < C \cdot \ell(\gamma)^2.$$

This contradicts the extremality of γ. ∎

We call a unit-speed curve $\gamma\colon S^1 \to X$ an $\mathcal{A}$-*extremal curve* if $I(\gamma) = C$.

Example 4.4. The particular shape of such extremal curves γ is only known in the following two situations.

(1) If $X = \mathbb{R}^n$ is Euclidean then, up to Euclidean motions, the extremal curves are given by the standard embedding of S^1 into $\mathbb{R}^2$. In particular, all such curves are planar and $C(X) = \frac{1}{4\pi}$ independently of $\mathcal{A}$. This follows from Reshetnyak's majorization theorem, [44], and the existence of a 1-Lipschitz retraction of X onto any of its linear subspaces.

(2) If X is a 2-dimensional normed space, then there is also less ambiguity in the choice of area functional $\mathcal{A}$. This is because all metric differentials $\mathrm{md}_p f$ of a Lipschitz map $f\colon D^2 \to X$ are either degenerate or give rise to normed spaces isometric to X. In particular, the shape of extremal curves does not depend on $\mathcal{A}$ and for area functionals $\mathcal{A}$ and $\bar{\mathcal{A}}$ one has

$$C^{\bar{\mathcal{A}}}(X) = \frac{\mathbf{J}^{\bar{\mathcal{A}}}(X)}{\mathbf{J}^{\mathcal{A}}(X)} \cdot C^{\mathcal{A}}(X). \tag{4.1}$$

Maybe somewhat surprisingly, the extremal curves γ do not correspond to the boundary contour of the unit ball B but rather to the boundary contour of the dual unit ball B^* under a suitable identification of X and X^*, see Section 4.4 in [48].

Although in these two examples the choice of $\mathcal{A}$ is immaterial for the shape of γ, it is very likely that this phenomenon is far from being true for a generic finite dimensional normed space X.

Remark 4.5. Lemma 1.6 does not hold for general Banach spaces X. Namely, by Theorem 1.2, if $C^{\mathcal{A}}(X) = \frac{1}{2\pi}$ and γ is an $\mathcal{A}$-extremal curve in X, then γ must be an isometric embedding of S^1. However, $C^{\mathrm{ht}}(\ell^1) = \frac{1}{2\pi}$ by Remark 4.11 below, and ℓ^1 does not admit such an isometric embedding by Remark 4.2.

4.3. Quadratic isoperimetric spectra

To prove Theorem 1.4, we fix $n \geq 2$ and endow $\mathbf{Ban}_n$ with the Banach–Mazur distance d_{BM}. It is given for $X, Y \in \mathbf{Ban}_n$ by

$$d_{\mathrm{BM}}(X, Y) := \inf\left\{\log(\|T\| \cdot \|T^{-1}\|) \mid T\colon X \to Y \text{ linear isomorphism}\right\}.$$

Endowed with the Banach–Mazur distance, $\mathbf{Ban}_n$ becomes a compact connected semi-metric space, see for example [48].

Lemma 4.6. $C^{\mathcal{A}}(\cdot)\colon \mathbf{Ban}_n \to \mathbb{R}$ *is continuous.*

Proof. Let $T\colon X \to Y$ be such that $\log(\|T\| \cdot \|T^{-1}\|) < \varepsilon$. Then for every $\gamma\colon S^1 \to X$ and $f\colon \bar{D}^2 \to X$ Lipschitz one has

$$\mathrm{e}^{-\varepsilon}\,\ell(\gamma) \leq \ell(T \circ \gamma) \leq \mathrm{e}^{\varepsilon}\,\ell(\gamma) \quad \text{and} \quad \mathrm{e}^{-2\varepsilon}\,\mathcal{A}(f) \leq \mathcal{A}(T \circ f) \leq \mathrm{e}^{2\varepsilon}\,\mathcal{A}(f). \tag{4.2}$$

From (4.2) it follows that

$$|\log(C^{\mathcal{A}}(X)) - \log(C^{\mathcal{A}}(Y))| < 8\varepsilon.$$

So $\log(C^{\mathcal{A}}(\cdot))$ is continuous on $\mathbf{Ban}_n$ and hence so is $C^{\mathcal{A}}(\cdot)$. ∎

At this point we prove the following variant of Theorem 1.4, which holds without assumptions on the area functional $\mathcal{A}$.

Theorem 4.7. $\mathrm{QIS}^{\mathcal{A}}(\mathbf{Ban}_n)$ *is a compact interval* $[l_n^{\mathcal{A}}, r_n^{\mathcal{A}}]$, *where*

$$\frac{1}{16} \leq l_n^{\mathcal{A}} \leq \frac{1}{4\pi} \leq r_n^{\mathcal{A}} < \frac{1}{2\pi}$$

and $r_n^{\mathcal{A}}$ *is nondecreasing in* n.

Proof. The quadratic isoperimetric spectrum $\mathrm{QIS}^{\mathcal{A}}(\mathbf{Ban}_n)$ is the image of the compact connected space $\mathbf{Ban}_n$ under the continuous map $C^{\mathcal{A}}(\cdot)$, and hence a compact interval $[l_n^{\mathcal{A}}, r_n^{\mathcal{A}}]$. By Example 4.4, we have $l_n^{\mathcal{A}} \leq \frac{1}{4\pi} \leq r_n^{\mathcal{A}}$. The inequality $l_n^{\mathcal{A}} \geq 1/16$ is implied by the discussion at the beginning of the proof of Lemma 4.3.

Fix $X \in \mathbf{Ban}_n$ such that $C^{\mathcal{A}}(X) = r_n^{\mathcal{A}}$ and an $\mathcal{A}$-extremal curve γ within X. By Lemma 1.7, γ cannot be an isometric embedding and hence Theorem 1.2 implies

$$r_n^{\mathcal{A}} = C^{\mathcal{A}}(\gamma) < \frac{1}{2\pi}\cdot$$

To see that $r_n^{\mathcal{A}}$ is nondecreasing it suffices to note that $C^{\mathcal{A}}(X \times \mathbb{R}) \geq C^{\mathcal{A}}(X)$. This is true because X is a 1-Lipschitz retract of $X \times \mathbb{R}$. ∎

We call $X \in \mathbf{Ban}_n$ an $\mathcal{A}$-*extremal space* if $C^{\mathcal{A}}(X) = r_n^{\mathcal{A}}$.

Example 4.8. Theorems 4.4.1 and 4.4.2 in [48] imply that any $X \in \mathbf{Ban}_2$ is $\mathcal{A}^{\mathrm{ht}}$ extremal. By comparison to $\mathcal{A}^{\mathrm{ht}}$ and equations (4.1), (2.5) and (2.6), the $\mathcal{A}$-quadratic isoperimetric spectra of $\mathbf{Ban}_2$ for $\mathcal{A} = \mathcal{A}^{\mathrm{ht}}, \mathcal{A}^{b}, \mathcal{A}^{m*}, \mathcal{A}^{\mathrm{ir}}$ are given as stated in Section 1.2. Except for $\mathcal{A}^{\mathrm{ht}}$, the (up to isometry) unique extremal space in all these situations is $\mathbb{R}^2_\infty$. By (2.5) we can also add

$$\mathrm{QIS}^{\mathrm{cr}}(\mathbf{Ban}_2) = \left[\frac{1}{16}, \frac{1}{4\pi}\right]$$

to the list, where by (2.4) the unique extremal space is the Euclidean plane.

For $n \geq 3$, the question which spaces are extremal remains completely open.

4.4. Lower bounds

To complete the proof of Theorem 1.4, by Theorem 4.7 it suffices to show that r_n^{ht} converges to $\frac{1}{2\pi}$ as $n \to \infty$. More precisely, we will prove (1.7). Remember that every separable metric space X admits an isometric embedding ι into the space l^∞ of bounded sequences endowed with the supremum norm. If X is compact, this Kuratowski embedding $\iota\colon X \to \ell^\infty$ is given by choosing a countable dense subset $\{x_1, x_2, x_3, \dots\}$ of X and setting

$$\iota(x) := (d(x, x_1), d(x, x_2), d(x, x_3), \dots).$$

Let $S_n := \{\phi_1, \dots, \phi_n\} \subset S^1$ be a cyclically ordered subset of equidistant points. The Kuratowski embedding gives an isometric embedding of S_n into $\mathbb{R}^n_\infty$. Hence the proof of (1.7), and thus the proof of Theorem 1.4, is completed by the following lemma.

Lemma 4.9. *Let X be a geodesic metric space, let $m \geq 2$ and let $e\colon S_m \to X$ be an isometric embedding. Then*

$$C^{\mathrm{ht}}(X) \geq \left(1 - \frac{4}{m}\right)\frac{1}{2\pi}\cdot$$

Proof. We may extend e to a 1-Lipschitz curve $\gamma\colon S^1 \to X$ by defining γ to equal a geodesic connecting $e(\phi_i)$ to $e(\phi_{i+1})$ on $[\phi_i, \phi_{i+1}]$. Let $\iota\colon S^1 \to l^\infty$ be the Kuratowski embedding. As l^∞ is an injective metric space, there is a 1-Lipschitz map $f\colon X \to l^\infty$ such that

$$f(\gamma(\phi_i)) = f(e(\phi_i)) = \iota(\phi_i)$$

for all $i = 1, \dots, m$, see for example [33]. Then, since $C^{\mathrm{ht}}(\ell^\infty) \leq \frac{1}{2\pi}$, Lemma 2.8 implies

$$(4.3) \qquad |\operatorname{Fill}^{\mathrm{ht}}(f \circ \gamma) - \operatorname{Fill}^{\mathrm{ht}}(\iota)| \leq \frac{m}{2\pi}\left(\frac{\ell(f \circ \gamma)}{m} + \frac{\ell(\iota)}{m}\right)^2 = \frac{m}{2\pi}\left(2\,\frac{2\pi}{m}\right)^2 = \frac{8\pi}{m}.$$

If g is a Lipschitz disc in X that bounds γ, then $f \circ g$ is a Lipschitz disc in ℓ^∞ that bounds $f \circ \gamma$ and satisfies $\mathcal{A}^{\mathrm{ht}}(f \circ g) \leq \mathcal{A}^{\mathrm{ht}}(g)$. From this observation together with (4.3) we deduce

$$\operatorname{Fill}^{\mathrm{ht}}(\gamma) \geq \operatorname{Fill}^{\mathrm{ht}}(f \circ \gamma) \geq \operatorname{Fill}^{\mathrm{ht}}(\iota) - \frac{8\pi}{m} = \left(1 - \frac{4}{m}\right) 2\pi.$$

As $\ell(\gamma) = 2\pi$, this implies the claim. ■

Remark 4.10. There are two observations that allow to push the lower bound on the constants $r_n^{\mathcal{A}}$ a bit further if one desires.

(1) The Kuratowski embedding of S_{2n} into $\mathbb{R}^{2n}_\infty$ carries more information than necessary. In fact one can forget about half of the coordinates and even obtain an isometric embedding of S_{2n} into $\mathbb{R}^n_\infty$. This leads to

$$(4.4) \qquad r_n^{\mathrm{ht}} \geq C^{\mathrm{ht}}(\mathbb{R}^n_\infty) \geq \left(1 - \frac{2}{n}\right)\frac{1}{2\pi}.$$

(2) If X is a polyhedral normed space such as $\mathbb{R}^n_\infty$, then by (2.4) and (2.6) one has $C^{\mathcal{A}}(X) > C^{\mathrm{ht}}(X)$ for $\mathcal{A} = \mathcal{A}^b, \mathcal{A}^{\mathrm{ir}}, \mathcal{A}^{m*}$. In particular, for all these area functionals the inequality (4.4) is even strict. Similarly one can obtain explicit upper bounds on the constants $C^{\mathcal{A}}(X)$ for $\mathcal{A} = \mathcal{A}^b, \mathcal{A}^{\mathrm{ht}}, \mathcal{A}^{\mathrm{cr}}$ and fixed polyhedral finite dimensional normed space X by comparing to $\mathcal{A}^{\mathrm{ir}}$ instead of $\mathcal{A}^{\mathrm{ht}}$.

Remark 4.11. There is also an isometric embedding j of S_n into $\mathbb{R}^n_1 = L^1(S_n)$ which is given by

$$(j(\phi))(\psi) = \begin{cases} \pi/n, & \text{if } (\phi, \psi, \tau(\phi)) \text{ is cyclically ordered and } \psi \neq \tau(\phi), \\ 0, & \text{otherwise.} \end{cases}$$

In particular, $C^{\mathrm{ht}}(\mathbb{R}^n_1) \geq \left(1 - \frac{4}{n}\right)\frac{1}{2\pi}$ and $C^{\mathrm{ht}}(\ell^1) = \frac{1}{2\pi}$.

We finish our paper with the proof of Theorem 1.3. More generally, we show that

$$(4.5) \qquad \operatorname{QIS}^{\mathcal{A}}(\mathbf{Ban}) = \{0\} \cup \left[\frac{1}{4\pi}, \frac{1}{2\pi}\right]$$

as soon as $\mathcal{A} \geq \mathcal{A}^{\mathrm{ht}}$.

Proof of (4.5). Let X be a Banach space of dimension ≥ 2. Theorem 1.2 in [16] implies that $C^{\mathcal{A}}(X) \leq \frac{1}{2\pi}$. Now let $V \subset X$ be a 2-dimensional linear subspace. Then Theorems 4.4.1 and 4.4.2 in [48] imply $C^{\mathrm{ht}}(V) = \frac{1}{4\pi}$. Thus our assumption $\mathcal{A} \geq \mathcal{A}^{\mathrm{ht}}$ and the remark after the statement of Theorem 3 in [11] give that

$$C^{\mathcal{A}}(X) \geq C^{\mathrm{ht}}(X) \geq C^{\mathrm{ht}}(V) = \frac{1}{4\pi}.$$

By Theorem 1.4 the interval $\left[\frac{1}{4\pi}, \frac{1}{2\pi}\right)$ is contained in $\operatorname{QIS}^{\mathcal{A}}(\mathbf{Ban})$. Thus the proof is completed by noting that $C^{\mathcal{A}}(\mathbb{R}) = 0$ and that by Remark 4.11 $C^{\mathcal{A}}(\ell^1) = \frac{1}{2\pi}$. ■

Acknowledgements. I would like to thank my PhD advisor Alexander Lytchak for great support in everything. Also I would like to thank Stefan Wenger for teaching an enlightening mini-course on "Dehn functions and large scale geometry" at the Young Geometric Group Theory winter school which took place in Les Diablerets in 2018. Furthermore, I would like to thank Vladimir Zolotov for giving me a short but very helpful explanation of optimal transport in dimension one. Finally, I want to thank the anonymous referees for their careful reading and the suggestions which helped a lot to improve the presentation of this article.

Funding. This work was partially supported by the DFG grant SPP 2026.

References

[1] Almgren, F. J. J.: Optimal isoperimetric inequalities. *Indiana Univ. Math. J.* **35** (1986), no. 3, 451–547.

[2] Álvarez Paiva, J. C. and Berck, G.: What is wrong with the Hausdorff measure in Finsler spaces. *Adv. Math.* **204** (2006), no. 2, 647–663.

[3] Álvarez Paiva, J. C. and Thompson, A. C.: Volumes on normed and Finsler spaces. In *A sampler of Riemann-Finsler geometry*, pp. 1–48. Math. Sci. Res. Inst. Publ. 50, Cambridge Univ. Press, Cambridge, 2004.

[4] Ambrosio, L. and Kirchheim, B.: Currents in metric spaces. *Acta Math.* **185** (2000), no. 1, 1–80.

[5] Bangert, V., Croke, C. B., Ivanov, S. V. and Katz, M. G.: Filling area conjecture and ovalless real hyperelliptic surfaces. *Geom. Funct. Anal.* **15** (2005), no. 3, 577–597.

[6] Benyamini, Y. and Lindenstrauss, J.: *Geometric nonlinear functional analysis. Vol. 1*. American Mathematical Society Colloquium Publications 48, American Mathematical Society, Providence, RI, 2000.

[7] Bernig, A.: Centroid bodies and the convexity of area functionals. *J. Differential Geom.* **98** (2014), no. 3, 357–373.

[8] Besicovitch, A. S.: On two problems of Loewner. *J. London Math. Soc.* **27** (1952), 141–144.

[9] Besson, G., Courtois, G. and Gallot, S.: Entropies et rigidités des espaces localement symétriques de courbure strictement négative. *Geom. Funct. Anal.* **5** (1995), no. 5, 731–799.

[10] Brady, N. and Bridson, M. R.: There is only one gap in the isoperimetric spectrum. *Geom. Funct. Anal.* **10** (2000), no. 5, 1053–1070.

[11] Burago, D. and Ivanov, S.: On asymptotic volume of Finsler tori, minimal surfaces in normed spaces, and symplectic filling volume. *Ann. of Math. (2)* **156** (2002), no. 2, 891–914.

[12] Burago, D. and Ivanov, S.: Boundary rigidity and filling volume minimality of metrics close to a flat one. *Ann. of Math. (2)* **171** (2010), no. 2, 1183–1211.

[13] Burago, D. and Ivanov, S.: Area minimizers and boundary rigidity of almost hyperbolic metrics. *Duke Math. J.* **162** (2013), no. 7, 1205–1248.

[14] Burago, D. and Ivanov, S.: Boundary distance, lens maps and entropy of geodesic flows of Finsler metrics. *Geom. Topol.* **20** (2016), no. 1, 469–490.

[15] Cabrelli, C. A. and Molter, U. M.: The Kantorovich metric for probability measures on the circle. *J. Comput. Appl. Math.* **57** (1995), no. 3, 345–361.

[16] Creutz, P.: Majorization by hemispheres and quadratic isoperimetric constants. *Trans. Amer. Math. Soc.* **373** (2020), no. 3, 1577–1596.

[17] Creutz, P.: Plateau's problem for singular curves. To appear in *Comm. Anal. Geom.*

[18] Creutz, P. and Fitzi, M.: The Plateau–Douglas problem for singular configurations and in general metric spaces. Preprint 2020, arXiv: 2008.08922.

[19] Creutz, P. and Romney, M.: The branch set of minimal disks in metric spaces. Preprint 2020, arXiv: 2008.07413.

[20] Creutz, P. and Soultanis, E.: Maximal metric surfaces and the Sobolev-to-Lipschitz property. *Calc. Var. Partial Differential Equations* **59** (2020), no. 5, Paper No. 177, 34 pp.

[21] Croke, C. B.: A sharp four-dimensional isoperimetric inequality. *Comment. Math. Helv.* **59** (1984), no. 2, 187–192.

[22] Fitzi, M. and Wenger, S.: Area minimizing surfaces of bounded genus in metric spaces. *J. Reine Angew. Math.* **770** (2021), 87–112.

[23] Gromov, M.: Filling Riemannian manifolds. *J. Differential Geom.* **18** (1983), no. 1, 1–147.

[24] Gromov, M.: Hyperbolic groups. In *Essays in group theory*, pp. 75–263. Math. Sci. Res. Inst. Publ. 8, Springer, New York, 1987.

[25] Gruber, P. M. and Schuster, F. E.: An arithmetic proof of John's ellipsoid theorem. *Arch. Math. (Basel)* **85** (2005), no. 1, 82–88.

[26] Holmes, R. D. and Thompson, A. C.: n-dimensional area and content in Minkowski spaces. *Pacific J. Math.* **85** (1979), no. 1, 77–110.

[27] Ivanov, S. V.: On two-dimensional minimal fillings. *Algebra i Analiz* **13** (2001), no. 1, 26–38.

[28] Ivanov, S. V.: Volumes and areas of Lipschitz metrics. *St. Petersburg Math. J.* **20** (2009), no. 3, 381–405.

[29] Ivanov, S. V.: Filling minimality of Finslerian 2-discs. *Proc. Steklov Inst. Math.* **273** (2011), no. 1, 176–190.

[30] Ivanov, S. V.: Local monotonicity of Riemannian and Finsler volume with respect to boundary distances. *Geom. Dedicata* **164** (2013), 83–96.

[31] Kirchheim, B.: Rectifiable metric spaces: local structure and regularity of the Hausdorff measure. *Proc. Amer. Math. Soc.* **121** (1994), no. 1, 113–123.

[32] Kleiner, B.: An isoperimetric comparison theorem. *Invent. Math.* **108** (1992), no. 1, 37–47.

[33] Lang, U.: Injective hulls of certain discrete metric spaces and groups. *J. Topol. Anal.* **5** (2013), no. 3, 297–331.

[34] Loeb, P. A. and Talvila, E.: Lusin's theorem and Bochner integration. *Sci. Math. Jpn.* **60** (2004), no. 1, 113–120.

[35] Lytchak, A. and Wenger, S.: Regularity of harmonic discs in spaces with quadratic isoperimetric inequality. *Calc. Var. Partial Differential Equations* **55** (2016), no. 4.

[36] Lytchak, A. and Wenger, S.: Area minimizing discs in metric spaces. *Arch. Ration. Mech. Anal.* **223** (2017), no. 3, 1123–1182.

[37] Lytchak, A. and Wenger, S.: Energy and area minimizers in metric spaces. *Adv. Calc. Var.* **10** (2017), no. 4, 407–421.

[38] Lytchak, A. and Wenger, S.: Intrinsic structure of minimal discs in metric spaces. *Geom. Topol.* **22** (2018), no. 1, 591–644.

[39] Lytchak, A. and Wenger, S.: Isoperimetric characterization of upper curvature bounds. *Acta Math.* **221** (2018), no. 1, 159–202.

[40] Lytchak, A., Wenger, S. and Young, R.: Dehn functions and Hölder extensions in asymptotic cones. *J. Reine Angew. Math.* **763** (2020), 79–109.

[41] Morgan, F. and Ritoré, M.: Isoperimetric regions in cones. *Trans. Amer. Math. Soc.* **354** (2002), no. 6, 2327–2339.

[42] Overath, P. and von der Mosel, H.: On minimal immersions in Finsler space. *Ann. Global Anal. Geom.* **48** (2015), no. 4, 397–422.

[43] Pu, P. M.: Some inequalities in certain nonorientable Riemannian manifolds. *Pacific J. Math.* **2** (1952), 55–71.

[44] Reshetnyak, Y. G.: Non-expansive maps in a space of curvature no greater than K. *Sibirsk. Mat. Zh.* **9** (1968), 918–927.

[45] Ritoré, M.: A proof by calibration of an isoperimetric inequality in the Heisenberg group $\mathbb{H}^n$. *Calc. Var. Partial Differential Equations* **44** (2012), no. 1-2, 47–60.

[46] Schulze, F.: Optimal isoperimetric inequalities for surfaces in any codimension in Cartan–Hadamard manifolds. *Geom. Funct. Anal.* **30** (2020), no. 1, 255–288.

[47] Stadler, S.: The structure of minimal surfaces in CAT(0) spaces. *J. Eur. Math. Soc. (JEMS)* **23** (2021), no. 11, 3521–3554.

[48] Thompson, A. C.: *Minkowski geometry*. Encyclopedia of Mathematics and its Applications 63, Cambridge University Press, 1996.

[49] Villani, C.: *Topics in optimal transportation*. Graduate Studies in Mathematics 58, American Mathematical Society, Providence, RI, 2003.

[50] Wenger, S.: Isoperimetric inequalities of Euclidean type in metric spaces. *Geom. Funct. Anal.* **15** (2005), no. 2, 534–554.

[51] Wenger, S.: Gromov hyperbolic spaces and the sharp isoperimetric constant. *Invent. Math.* **171** (2008), no. 1, 227–255.

[52] Wenger, S.: Spaces with almost Euclidean Dehn function. *Math. Ann.* **373** (2019), no. 3-4, 1177–1210.

Received April 10, 2020; revised April 8, 2021. Published online November 2, 2021.

Paul Creutz
Mathematisches Institut der Universität zu Köln,Weyertal 86-90, 50931 Köln, Germany;
Paul.Creutz@gmx.de