ON TWO-DIMENSIONAL MINIMAL FILLINGS

S. V. Ivanov

Abstract. We consider Riemannian metrics in the two-dimensional disk D (with boundary). We prove that, if a metric g_0 is such that every two interior points of D are connected by a unique geodesic of g_0, or if g_0 can be extended to a complete metric without conjugate points in $\mathbf{R}^2$, then the Riemannian area of g_0 is not greater than the area of any other metric g in which the distances between boundary points of D are not less than those in g_0. Previously this fact was known only in the case when g_0 has constant curvature. We give a generalization of the main result to the Finslerian case and an interpretation of it in terms of simply connected Lipschitz surfaces with a fixed boundary in a Banach space.

Introduction

Let g_0 be a Riemannian metric in the two-dimensional disc D (with boundary). We say that g_0 is *area-minimizing* if for any Riemannian metric g in D such that

$$\operatorname{dist}_g(x,y) \geq \operatorname{dist}_{g_0}(x,y)$$

for all $x, y \in \partial D$, one has

$$\operatorname{Area}(D,g) \geq \operatorname{Area}(D,g_0).$$

Here dist_g denotes the Riemannian distance in (D,g) and Area is the Riemannian area.

The main purpose of this paper is to prove the following

0.1. Main Theorem. *Let g_0 be a Riemannian metric in the two-dimensional disc such that every two points in the disc's interior are connected by a unique geodesic of g_0. Then g_0 is area-minimizing.*

Previously known results. The statement of Theorem 0.1 is known in the cases when g_0 is a metric of constant curvature K, cf. [G, §5.5]. If $K = 0$, it follows from Besicovitch inequality ([Be], see also [BuZ]). If $K = -1$, it follows from the volume entropy inequality for hyperbolic manifolds ([K], for higher-dimensional case see [BCG]). If $K = 1$, one may assume that (D, g_0) is isometric to the unit hemisphere. In this case, the statement is equivalent to the theorem of Pu [P] about the isosystolic constant of $\mathbf{RP}^2$.

Typeset by $\mathcal{AMS}$-TeX

Metrics without conjugate points in $\mathbf{R}^2$. The condition on g_0 is equivalent to that the geodesics in the interior of the disc have no conjugate points and the boundary of the disc contains no concave arcs. In particular, if g_0 is a complete Riemannian metric without conjugate points in $\mathbf{R}^2$, its restriction to any g_0-convex domain $D \subset \mathbf{R}^2$ satisfies this condition. It is easy to see that any bounded set in such a space $(\mathbf{R}^2, g_0)$ is contained in a bounded convex domain. Since sub-domains of an area-minimizing domain obviously inherit the area-minimizing property, we obtain the following

0.2. Corollary. *Let g_0 be a complete Riemannian metric without conjugate points in $\mathbf{R}^2$ and let D be a compact region (not necessarily homeomorphic to the disc) with a piecewise smooth boundary. Then the restriction of g_0 on D is area-minimizing.*

Finslerian case. Theorem 0.1 remains valid for Finsler metrics if the Finsler area is defined as the two-dimensional Holmes–Thompson volume (i. e. the symplectic volume of the unit co-tangent bundle). To avoid the discussion of Finsler volumes, We do not formulate the Finsler version of the theorem in this introduction. The necessary definitions and modifications for the proof are given in §3.

Relation to filling volumes. The area-minimizing property is related to the notion of *filling volume* introduced by Gromov [G]. The *filling volume* of a manifold equipped with a distance function is the greatest lower bound for the volume of a (Finsler) space spanning the given manifold as a boundary and inducing on it a distance function which is not less than the given one. One can vary this definition by choosing a special class of spaces within which the volume is to be minimized. For example, one may restrict this class to a fixed topological type. In this context, Theorem 0.1 can be formulated as follows:

(D, g_0) is a "minimal filling" of its own boundary in the class of spaces homeomorphic to the 2-disc.

It remains unclear whether the assumption "homeomorphic to the 2-disc" can be dropped, that is, whether it is possible to replace (D, g) by a (D', g) where D' is a 2-manifold with $\partial D' = \partial D$. This is not known even in the case of the theorem of Pu, i. e. when (D, g_0) is the unit hemisphere.

Remarks on the equality case. It is natural to conjecture that the equality $\mathrm{Area}(D, g) = \mathrm{Area}(D, g_0)$ implies that (D, g) and (D, g_0) are isometric. This can be proved in a number of partial cases, for example, if g_0 is nonpositively curved. Indeed, one can see from the proof of Theorem 0.1 that the equality of areas implies that $\mathrm{dist}_g|_{\partial D} = \mathrm{dist}_{g_0}|_{\partial D}$. Then the question reduces to the boundary rigidity problem, see [C1] and references therein. The boundary rigidity for nonpositively curved metrics is proved in [C2].

In the Finsler case the equality $\mathrm{Area}(D, g) = \mathrm{Area}(D, g_0)$ does not imply that the spaces are isometric. For every metric g_0 satisfying the conditions of Theorem 0.1, there is an infinite-parameter family of non-isometric Finsler metrics g with $\mathrm{dist}_g|_{\partial D} = \mathrm{dist}_{g_0}|_{\partial D}$ and $\mathrm{Area}(D, g) = \mathrm{Area}(D, g_0)$. In fact, any C^2-small perturbation of the function $\mathrm{dist}_{g_0}|_{D \times \partial D}$ fixed at $\partial D \times \partial D$ is a distance function of a Finsler metric having the same area as g_0.

Structure of the paper. The proof of Theorem 0.1 is contained in sections 1 and 2. In sections 3 and 4 we briefly discuss the Finsler case and related issues about minimal surfaces in Banach spaces.

The proof (more precisely, the arguments in sections 2 and 3) is based on the same ideas as a similar result about asymptotic areas of periodic Finsler metrics obtained jointly with D. Burago (see the forthcoming paper [BuI]). Section 1 contains a construction that allows us to apply the technique from [BuI] to the discussed problem.

1. Special distance-like functions

Note that the assumption on g_0 implies that every geodesic in (D, g_0) is a shortest path between any two its points. Furthermore, any boundary point can be jointed to any interior point by a unique g_0-geodesic. We denote by $|xy|_0$ and $|xy|$ the Riemannian distances between points x and y in the spaces (D, g_0) and (D, g) respectively.

For a $p \in \partial D$ define a function $f_p : D \to \mathbf{R}$ by

$$f_p(x) = \max_{q \in \partial D}(|pq|_0 - |xq|). \tag{1}$$

The maximum here is achieved due to compactness of ∂D.

1.1. Lemma. *Let $p \in \partial D$. Then*

1. *f_p is a nonexpanding function with respect to g, i. e. $|f_p(x) - f_p(y)| \le |xy|$ for all $x, y \in D$.*
2. *If $x \in \partial D$, then $f_p(x) = |px|_0$.*
3. *If $g = g_0$, then $f_p(x) = |xp|$ for all $x \in D$, $p \in \partial D$.*

Proof. 1. f_p is a supremum of nonexpanding functions $x \mapsto -|xq| + |pq|_0$, hence it is nonexpanding.

2. Since $x \in \partial D$, we have $|xq| \ge |xq|_0$ for all $q \in \partial D$. Then $|pq|_0 - |xq| \le |pq|_0 - |xq|_0 \le |px|_0$ by the triangle inequality. One the other hand, letting $q = x$ yields $|pq|_0 - |xq| = |pq|_0$.

3. Suppose $g = g_0$ and let $q\partial D$ be a point where the geodesic from p through x hits the boundary. Since this geodesic is minimal, we have $|pq| = |px| + |xq|$. Hence $f_p(x) \ge |pq| - |qx| = |px|$. On the other hand, $f_p(x) \le f_p(p) + |px| = |px|$ by the first two statements of the lemma. $\square$

Remark. Let $\mathcal{X} = \ell_\infty(\partial D)$, the Banach space of bounded functions $\varphi\colon \partial D \to \mathbf{R}$ with the norm $\|\varphi\|_\infty = \sup|\varphi|$. Consider the map $\mathcal{F}\colon D \to \mathcal{X}$ defined by $\mathcal{F}(x)(p) = f_p(x)$. The first statement of Lemma 1.1 means that $\|\mathcal{F}(x) - \mathcal{F}(y)\|_\infty \le |xy|$, i. e. $\mathcal{F}$ is a nonexpanding map. The second statement implies that the restriction of $\mathcal{F}$ to ∂D does not depend on g and coincide with the canonical Gromov's embedding of the metric space $(\partial D, \mathrm{dist}_{g_0})$ into $\mathcal{X}$. There are other ways to obtain a map with these properties (for example, one could define $f_p(x) = \min_{q\in\partial D}(|pq|_0 + |qx|)$). However the maps defined by (1) satisfy an additional requirement on their derivatives (Lemma 1.2) which is essential for our proof.

We denote by $\operatorname{grad} f$ the gradient of a function $f\colon D \to \mathbf{R}$ with respect to g. Since every function f_p is Lipschitz, $\operatorname{grad} f_p$ is defined almost everywhere and is a measurable vector field.

We use the notation $|\cdot|$ for the fiber-wise Euclidean norms on TD and T^*D determined by g. The bundles of unit vectors and co-vectors are denoted by UTD and UT^*D, respectively. Fix an orientation of D. It induces cyclic orderings on ∂D and every fiber of UTD. By means of the natural isomorphism $TD \cong T^*D$ a cyclic order is then defined on the fibers of UT^*D.

1.2. Lemma. 1. *If $p \in \partial D$, x is an interior point of D and f_p is differentiable at $x \in D$, then $|df(x)| = 1$.*

2. *Let x be an interior point of D and $\{p_i\}_{i=1}^n$ be a collection of points in ∂D. Suppose that the functions f_{p_i} are differentiable at x and their derivatives $df_{p_i}(x)$ are mutually different. Then the cyclic order of these derivatives in UT_x^*D is the same as the cyclic order of the points p_i in ∂D.*

Proof. We will prove the statements for gradients instead of derivatives and will use Lemma 1.1 without referring to it explicitly. Fix a point x in the interior of D. We call a $q \in \partial D$ a *point of maximum for a* $p \in \partial D$ if maximum in (1) is achieved at q, i. e. $f_p(x) = |pq|_0 - |xq|$. Let $p \in \partial D$, f_p be differentiable at x, q be a point of maximum for p, and $\gamma\colon [0, |xq|] \to D$ be a unit-speed g-shortest curve connecting x to q, i. e. $\gamma(0) = x$, $\gamma(|xq|) = q$ and $|\gamma(t)\gamma(t')| = |t - t'|$ for all $t, t' \in [0, |xq|]$. An initial arc of γ is contained in the interior of D and hence is a geodesic. In particular, γ is differentiable at 0 and $|\gamma'(0)| = 1$. Since f_p is a nonexpanding function and $f_p(q) = |pq|_0 = f_p(x) + |xq|$, we have $f_p(\gamma(t)) = f_p(x) + t$ for all $t \in [0, |xq|]$. In other words, f_p grows at unit rate along γ. Since this is a maximal possible growth rate for f_p, it follows that $\operatorname{grad} f_p(x) = \gamma'(0)$. In particular, $|\operatorname{grad} f_p(x)| = 1$.

Before proving the second statement of the lemma, observe the following: under the same assumptions as in the first one, a nearest to x point of maximum for p is unique and is connected to x by a unique minimal g-geodesic. Indeed, if q is a nearest point of maximum and γ is as above, then $\gamma(t) \notin \partial D$ for all $t \in [0, |xq|)$ because otherwise the equality $f_p(\gamma(t)) = f_p(x) + t = f_p(x) + |x\gamma(t)|$ implies that $\gamma(t)$ is another point of maximum. Thus γ (except the endpoint) is contained in the interior of D, hence it coincides with the unique g-geodesic emanating from x with the initial velocity $\operatorname{grad} f_p(x)$. The uniqueness of q now follows from the fact that q is the point where this geodesic hits ∂D.

It is sufficient to prove the second statement of the lemma for $n = 3$ only, because the cyclic ordering of a set is determined by orderings of its three-element subsets. Let $\{p_i\}_{i=1}^3$ be as in the second statement of the lemma. For each $i \in \{1, 2, 3\}$ let q_i be the nearest to x point of maximum for p_i, and let γ_i be the (unique) minimal geodesic connecting x to q_i. Observe that the points q_i are mutually different because the vectors $\gamma_i'(0) = \operatorname{grad} f_{p_i}(x)$ are. Since D is a two-dimensional disc, the Jordan curve theorem implies that the points q_i obey the same cyclic order in ∂D as the vectors $\gamma_i'(0)$ do in UT_xD. Therefore it is sufficient to prove that the cyclic orderings of the triples (p_1, p_2, p_3) and (q_1, q_2, q_3) in ∂D coincide. As the first step, we prove the following

1.3. Lemma. $q_i \neq p_i$ *for all* $i \in \{1, 2, 3\}$. *Furthermore, if* $i, j \in \{1, 2, 3\}$ *and* $i \neq j$, *the pair* $\{p_i, q_j\}$ *does not separate* $\{p_j, q_i\}$ *in* ∂D *provided that the four mentioned points are mutually different.*

(We say that a pair $\{a, b\}$ of points in ∂D separates a pair $\{c, d\}$ if c and d belong to different components of the complement $\partial D \setminus \{a, b\}$).

Proof. Suppose the contrary. For definiteness, we may assume that $p_1 = q_1$ or $\{p_1, q_2\}$ separate $\{p_2, q_1\}$ in ∂D. In both cases, any curve in D connecting p_1 to q_2 intersect (possibly at an endpoint) any curve connecting p_2 to q_1. Let z be a common point of two g_0-shortest paths connecting p_1 to q_2 and p_2 to q_1. Then

$$|p_1q_2|_0 + |p_2q_1|_0 = |p_1z|_0 + |zq_2|_0 + |p_2z|_0 + |zq_1|_0 \geq |p_1q_1|_0 + |p_2q_2|_0$$

by the triangle inequality. Hence

$$|p_1q_2|_0 - |xq_2| + |p_2q_1|_0 - |xq_1| \geq |p_1q_1|_0 - |xq_1| + |p_2q_2|_0 - |xq_2| = f_{p_1}(x) + f_{p_2}(x).$$

On the other hand, $|p_1q_2|_0 - |xq_2| \leq f_{p_1}(x)$ and $|p_2q_1|_0 - |xq_1| \leq f_{p_2}(x)$ by the definition of f_{p_i}. Therefore these inequalities turn to equalities, in other words, both q_1 and q_2 are points of maximum for p_1 as well as for p_2. This contradicts the choice of q_i and the uniqueness of the nearest point of maximum. □

Lemma 1.3 alone implies the desired coincidence of cyclic orders. We will now prove this implication (which is just a combinatorial fact about six points in a circle and could be verified by exhaustion of possible cyclic orders). We may assume that the six points $\{p_i\}$ and $\{q_i\}$ are mutually different. This can be achieved by a small perturbation (it is easy to see that a small perturbation does not invalidate the property formulated in Lemma 1.3). Consider two cases.

Case 1. For all pairs (i, j) where $i, j \in \{1, 2, 3\}$ and $i \neq j$, the points p_i and q_i separate $\{p_j, q_j\}$ in ∂D. In this case there is a homeomorphism $\varphi\colon \partial D \to S^1$ which maps each pair (p_i, q_i) to a pair of opposite points in the circle. Since the central symmetry of S^1 preserves the orientation, the triples $\{\varphi(p_i)\}_{i=1}^3$ and $\{\varphi(q_i)\}_{i=1}^3$ are ordered similarly. So are the triples (p_1, p_2, p_3) and (q_1, q_2, q_3).

Case 2. For definiteness, $\{p_1, q_1\}$ do not separate $\{p_2, q_2\}$. By Lemma 1.3, $\{p_1, q_2\}$ cannot separate $\{p_2, q_1\}$. Hence $\{p_1, p_2\}$ separates $\{q_1, q_2\}$. The points p_1, q_1, p_2 and q_2 divide ∂D into four arcs: $[p_1, q_1]$, $[q_1, p_2]$, $[p_2, q_2]$ and $[q_2, p_1]$, denoted according to their endpoints. We now consider two subcases: $p_3 \in [p_1, q_1]$ and $p_3 \in [q_1, p_2]$. Other possible locations of p_3 can be reduced to these two by interchanging the indices 1 and 2. We will show that $q_3 \in [q_1, p_2] \cup [p_2, q_2]$ in both subcases.

Subcase 2a. $p_3 \in [p_1, q_1]$. Then p_3 divides $[p_1, q_1]$ into two arcs: $[p_1, p_3]$ and $[p_3, q_1]$. Applying Lemma 1.3 to $i = 3$ and $j = 1$, we obtain that $q_3 \notin [p_3, q_1]$. Applying Lemma 1.3 to $i = 3$ and $j = 2$, we obtain that $q_3 \notin [q_2, p_1] \cup [p_1, p_3]$. This determines the position of q_3 with respect to q_1 and q_2, namely $q_3 \in [q_1, p_2] \cup [p_2, q_2]$.

Subcase 2b. $p_3 \in [q_1, p_2]$. Then p_3 divides $[q_1, p_2]$ into two arcs: $[q_1, p_3]$ and $[p_3, p_2]$. Applying Lemma 1.3 to $i = 3$ and $j = 2$, we obtain that $q_2 \in [p_3, p_2] \cup [p_2, q_2] \subset [q_1, p_2] \cup [p_2, q_2]$.

Since $q_3 \in [q_1, p_2] \cup [p_2, q_2]$, it follows that $(q_1, q_3, q_2) \sim (q_1, p_2, q_2) \sim (p_1, q_1, p_2) \sim (p_1, p_3, p_2)$ in both subcases, where $\sim$ denotes the coincidence of cyclic orders. This completes the proof of Lemma 1.2. □

2. Proof of Theorem 0.1

Let $P = \{p_1, \dots, p_n\}$ be a cyclically ordered collection of points in ∂D. Fix a homeomorphism $\varphi\colon \partial D \to S^1$ and let $\delta = \delta(P)$ denote the maximum length of the arcs into which the points $\varphi(p_i)$ divide the circle. The reader should think of P as a member of a family (or a sequence) of partitions with $\delta(P) \to 0$. All indices below are taken modulo $n = n(P)$.

Define a 2-form ω_P in D by

$$\omega_P = \sum_{i=1}^{n} df_{p_i} \wedge df_{p_{i+1}}$$

where f_{p_i} are functions defined in the previous section. Since f_{p_i} are Lipschitz functions, ω_P is an almost everywhere defined measurable differential form.

Let dA denote the oriented area form of the metric g. Every measurable 2-form in D is represented in a form $\omega = A(\omega)\, dA$ where $A(\omega)$ is a measurable real-valued function on D.

2.1. Lemma. *$A(\omega_P) \le 2\pi$ almost everywhere. If $g = g_0$, then $A(\omega_P) \to 2\pi$ a. e. as $\delta(P) \to 0$.*

Proof. 1. Fix a point x in the interior of D such that the functions f_{p_i} are differentiable at x. We will show that $A(\omega_P)(x) \le 2\pi$. Denote $df_{p_i}(x)$ by v_i. Observe that the quantity $\frac{1}{2}A(df_{p_i} \wedge df_{p_{i+1}})$ equals the oriented area of the triangle $\triangle 0 v_i v_{i+1}$ in $T^*_x D$ (the Euclidean structure in $T^*_x D$ is determined by g). To estimate the sum of these quantities we will study how these triangles can overlap.

We may assume that $v_i \ne v_{i+1}$ for all i. Indeed, if $v_i = v_{i+1}$, the point p_i can be removed from the collection P without affecting the value of ω_P at the discussed point x. If the co-vectors v_i are mutually different, Lemma 1.2 implies they are cyclically ordered in $UT^*_x D$. Therefore the sum of oriented areas of the triangles $\triangle 0 v_i v_{i+1}$ equals the area of the convex polygon $v_1 v_2 \dots v_n$ in $T^*_x D$. Since this polygon is inscribed in the unit circle $UT^*_x D$, is area is less than π. Hence

$$A(\omega_P)(x) = 2 \sum_i (\text{oriented area of } \triangle 0 v_i v_{i+1}) < 2\pi.$$

If some of the vectors v_i coincide, we may assume that $v_1 = v_k$ for some k, $2 < k < n$. Consider $i \in \{2, \dots, k-1\}$ and $j \in \{k+1, \dots, n\}$ such that $v_i \ne v_1$ and $v_j \ne v_1$. Since the cyclic orders of the triples (p_i, p_j, p_1) and (p_i, p_j, p_k) are different, Lemma 1.2 implies that $v_i = v_j$. It follows that all vectors v_i that are not equal to v_1, coincide, i. e. there are only two different vectors among $\{v_i\}_{i=1}^n$. In this case it is easy to see that $A(\omega_P)(x) = 0$.

2. Let $g = g_0$. Then $f_p(x) = |px|$ by Lemma 1.1, hence f_p is differentiable everywhere in the interior of D and $\operatorname{grad} f_p(x)$ is the vector opposite to the initial velocity of the unique geodesic from x to p. Therefore the rule $p \mapsto df_p(x)$ defines a homeomorphism from ∂D to $UT^*_x D$. The vertices v_i of the polygon discussed above are images of the points p_i under this homeomorphism. Hence this polygon approaches the circle $UT^*_x D$ and its area approaches π as $\delta(P) \to 0$. Thus $A(\omega_P)(x) \to 2\pi$ as $\delta(P) \to 0$ for every interior point x. $\square$

Proof of Theorem 0.1. The 2-form ω_P is a pull-back of the form $\sum dx_i \wedge dx_{i+1}$ in $\mathbf{R}^n$ under the Lipschitz map $F_P \colon D \to \mathbf{R}^n$ defined by

$$F_P(x) = (f_{p_1}(x), \dots, f_{p_n}(x)).$$

The second statement of Lemma 1.1 implies that the restriction $F_P|_{\partial D}$ is determined by P and g_0. Since the form $\sum dx_i \wedge dx_{i+1}$ is closed, it follows that the integral $\int_D \omega_P$ depend only on P and g_0 but not on g. The first part of Lemma 2.1 implies that

$$I(P) := \int_D \omega_P \le 2\pi \operatorname{Area}(D, g).$$

On the other hand, the second part of Lemma 2.1 implies that

$$\lim_{\delta(P) \to 0} I(P) = 2\pi \operatorname{Area}(D, g_0)$$

and the theorem follows. $\square$

Remark. The integral $\int_D \omega_P$ can be explicitly rewritten in terms of $\operatorname{dist}_{g_0}|_{\partial D}$ by means of the Stokes' formula. The resulting expression is a finite-sum approximation of the Santalo's integral formula for the area of a metric with no conjugate points in terms of boundary distances, cf. [S], [G].

3. The Finsler case

A *Finsler manifold* is a smooth manifold M equipped with a function $\Phi\colon TM \to \mathbf{R}$ whose restriction on every fiber T_xM is a vector-space norm. Riemannian manifolds are a special case where these restrictions are Euclidean norms. Similarly to the Riemannian case, one defines the length of smooth curves and the distance between points. We will mainly consider smooth (outside the zero section) and strictly convex Finsler structures. The latter means that the function Φ^2 has positive definite second derivatives on the fiber $T_xM \setminus \{0\}$ for every $x \in M$. These requirements ensure the existence of smooth geodesics and an exponential map.

The co-tangent bundle of a Finsler manifold (M^n, Φ) is naturally equipped with a fiber-wise norm dual to Φ. We use the notation $|\cdot|$ for both Φ and the dual norm.

The *Holmes–Thompson volume* of (M, Φ) equals, by definition, the canonical symplectic volume of the set $B^*M := \{w \in T^*M : |w| \leq 1\}$, divided by the volume of the Euclidean unit ball in $\mathbf{R}^n$. For an oriented manifold, this can be interpreted as follows. Every fiber T^*_xM as a vector space carries a natural n-dimensional measure valued in $\bigwedge^n T^*_xM$. Thus the measures of the balls $B^*_xM = B^*M \cap T^*_xM$, $x \in M$, define a section of $\bigwedge^n T^*M$, that is, a differential n-form on M. The integration of this form over M yields the canonical symplectic volume of the set B^*M.

Unlike in the Riemannian case, there are many different notions of Finsler volume suitable for different purposes. For example, in the Finsler case the Hausdorff measure is generally not equal to the Holmes–Thompson volume.

Theorem 0.1 remains valid if g_0 and g are smooth and strictly convex Finsler structures and Area means the two-dimensional Homes–Thompson volume. Note that the regularity assumptions on g can be relaxed by means of an approximation argument.

The proof of the theorem works for the Finsler case with little modifications. The only thing to change in the proof is to get rid of Riemannian gradients. Our use of gradients was twofold. First, in Lemma 1.2 we use the fact that the gradient is a vector pointing to the direction of the fastest growth of a function, and that $|\operatorname{grad} f| = |df|$. Second, we use that the gradients are obtained from the derivatives by means of a fiber-wise orientation-preserving homeomorphism from T^*D to TD (in the Riemannian case, this homeomorphism is a linear map). These properties hold in the Finsler case as well if one defines the "gradient" of a function as the Legendre transform of its derivative with respect to Φ^2 where Φ is the Finsler structure. With this definition of "Finsler gradient", the arguments in the proof of Theorem 0.1 lead to estimates of the Holmes–Thompson area (compare the proof of Lemma 2.1 with the above interpretation of the Finsler volume as the integral of a differential form defined by the sets B^*_xM).

Relations to minimal surfaces

Finally we point out an interpretation of Theorem 0.1 and its proof in terms of minimal surfaces in Banach spaces. One can define the area of a two-dimensional Lipschitz surface in a Banach space as the two-dimensional Holmes–Thompson volume of its intrinsic metric. This intrinsic metric corresponds to a measurable Finsler structure and the definition of volume trivially extends to this case. We will use the loose notation $\mathrm{Area}(F(D))$ for the area of Lipschitz map F from D to a Banach space,

Let $\mathcal{X} = \ell_\infty(\partial D)$. Consider the map

$$F_0 \colon D \to \mathcal{X}, \quad F_0(x)(p) = |xp|_0; \quad x \in D,\ p \in \partial D.$$

The conditions on g_0 imply that F_0 is a distance-preserving map from (D, g_0) to $\mathcal{X}$. Regardless to the choice of a Finsler area functional, the statement of Theorem 0.1 is equivalent to that the resulting surface $F_0(D) \subset \mathcal{X}$ is area-minimizing in the class of Lipschitz surfaces parameterized by the two-dimensional disc and spanning the same boundary.

One implication is almost obvious. Let $F \colon D \to \mathcal{X}$ be a Lipschitz map with $F|_{\partial D} = F_0|_{\partial D}$ and let dist_F denote the intrinsic distance induced by F. Then for all $p, q \in \partial D$

$$\mathrm{dist}_F(p,q) \geq \|F(p) - F(q)\| = \mathrm{dist}_{g_0}(p,q).$$

By Theorem 0.1 extended to the case where g is a measurable Finsler structure, it follows that

$$\mathrm{Area}(F(D)) \geq \mathrm{Area}(D, g) = \mathrm{Area}(F_0(D)).$$

The converse implication is seen from the proof of the theorem. As a matter of fact, we proved just the inequality between areas of surfaces: $\mathrm{Area}(F_0(D)) \leq \mathrm{Area}(F(D))$, where $F \colon D \to \mathcal{X}$ is a map defined by means of the functions f_p, see the remark after the proof of Lemma 1.1. The theorem itself is an immediate consequence of this inequality, because F_0 is distance-preserving w. r. t. g_0 and F is distance-nonincreasing w. r. t. g. Note that this argument does not depend on the choice of a Finsler area functional, provided that the area depends monotonously on the metric.

The proof of the inequality $\mathrm{Area}(F_0(D)) \leq \mathrm{Area}(F(D))$ for the case of Holmes–Thompson area is contained in Lemma 2.1 and uses a method similar to the calibrating form technique. Namely we constructed a closed 2-form ($\sum dx_i \wedge dx_{i+1}$) which does not exceed the area form at $F(D)$ and equals (more precisely, almost equals) the area form at the surface $F_0(D)$ (more precisely, its finite-dimensional approximation). However the comparison of this 2-form ω_P and the area form at $F(D)$ essentially relies on the fact (Lemma 1.2) that the tangent planes of finite-dimensional approximations of the surface belong to a special set of planes, namely those planes in which the restrictions of coordinate functions are cyclically ordered. (This fixes a combinatorial type of the intersection of the plane with the coordinate cube.) It can be shown that a true calibrating form for $F_0(D)$ does not exist. Moreover, the surface $F_0(D)$ is not area-minimizing in the class of rational chains, cf. [BuI] for similar examples.

References

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

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

[BuI] D. Burago and S. Ivanov, *On asymptotic volume of Finsler tori, minimal surfaces in Banach spaces, and symplectic filling volume*, in preparation.

[BuZ] Yu. Burago and V. Zalgaller, *Geometric inequalities*, Springer-Verlag, 1988.

[C1] C. B. Croke, *Rigidity and the distance between boundary points*, J. Differential Geometry **33** (1991), 445–464.

[C2] C. B. Croke, *Rigidity for surfaces of non-positive curvature*, Comment. Math. Helv. **65** (1990), 150–169.

[G] M. Gromov, *Filling Riemannian manifolds*, J. Diff. Geom. **18** (1983), 1–147.

[K] A. Katok, *Entropy and closed geodesics*, Technical report, University of Maryland (1981).

[P] P. Pu, *Some inequalities in certain non-orientable Riemannian manifolds*, Pacific J. Math. **2** (1952), 55–71.

[S] L. A. Santaló, *Integral geometry and geometric probability,* Encyclopedia Math. Appl., Addison-Wesley, London, 1976.

Steklov Institute of Mathematics at St. Petersburg
E-mail address: svivanov@pdmi.ras.ru