

APOLLONIAN ISOMETRIES OF REGULAR DOMAINS ARE MÖBIUS MAPPINGS

PETER HÄSTÖ AND ZAIR IBRAGIMOV

ABSTRACT. The Apollonian metric is a generalization of the hyperbolic metric, defined in a much larger class of open sets. However, since it was introduced by Beardon in 1998, it has remained an open question what its isometries are. Beardon first raised this question and asked if the Apollonian isometries were just Möbius mappings. In this paper we show that this is the case in open sets with regular, for instance C^1 , boundary.

1. INTRODUCTION

The Apollonian metric α_D is a generalization of the hyperbolic metric which is defined in arbitrary open sets D of \mathbb{R}^n . It was introduced by Beardon in 1998 [3], but it later turned out that the same metric had been studied previously by Barbilian [1] (see [18], and for some further developments, [6, 7]). A very popular generalization of the hyperbolic metric in higher dimensions is the quasihyperbolic metric [10]. The Apollonian metric differs from this metric by equaling the hyperbolic metric in the ball and being Möbius invariant (rather than quasi-invariant). Another nice feature is that the Apollonian metric is very simple to evaluate.

In the 1998 paper Beardon asked if the isometries of the Apollonian metric are just Möbius mappings. In this paper we show that this is the case in open sets with regular boundary. Although we are not able to treat the completely general case, we would like to emphasize that previously such a result was known only for complements of sets of constant width.

Let us start by reviewing what has been done on the isometry question. Consider a planar domain the boundary of which is a compact subset of the extended negative real axis. In this setting Beardon proved that a conformal Apollonian isometry is a Möbius mapping, [3, Theorem 1.3]. In [16, Theorem 1] Ibragimov derived the same conclusion without the conformality assumption.

A different approach was used by Gehring & Hag [9]. Their proof uses geodesics of the Apollonian metric to show that every Apollonian isometry of a disk is a Möbius mapping [9, Theorem 3.29]. Unfortunately, the Apollonian metric does not in general have many geodesics (more on this later). Ibragimov extended the method of geodesics to its limits in [15], studying complements of sets of constant width. In this case there are still lots of geodesics through a single distinguished point. So he could show that in this case all Apollonian isometries are Möbius mappings.

Since we can get no further with geodesics we introduce the notion of a pseudogeodesic line in this paper (Definition 4.5). These are curves which share some

2000 *Mathematics Subject Classification.* 30F45 (primary), 30C65 (secondary).

The first author was supported in part by the Finnish Academy of Science and Letters.

crucial features with geodesics. On the other hand, it turns out that there exists at least one such pseudogeodesic line through every point (Theorem 4.6). This is one of our main tools.

The second new element in our approach is that we start by looking at what Apollonian isometries do to the boundary of the domain. Using estimates of the metric we show in Section 5 that the distortion of an Apollonian isometry decreases toward the boundary, and thus it has to act as a Möbius mapping on the boundary of the domain. Notice that this second result holds for all domains, not only regular ones. Hence it will probably be useful in solving the general case of the Apollonian isometry problem.

Combining these elements we prove our main result (Theorem 6.4), that every Apollonian isometry of a regular domain is a Möbius mapping. We start by summarizing the notation that we will use (Section 2) and by defining the Apollonian metric and giving some of its properties (Section 3).

Although this article will deal only with the problem of isometries, we conclude the introduction by mentioning some other results regarding this metric, by way of motivation. Beardon [3] and Rhodes [19] derived comparison results with the hyperbolic metric in convex planar domains. Gehring & Hag [9] showed that the Apollonian metric is comparable to the hyperbolic metric in a simply connected planar domain if and only if the domain is a quasidisk. Ibragimov [17] and Hästö [11] studied the relations between the Apollonian and quasihyperbolic metrics. They showed, in particular, that the two metrics are equivalent if the domain is a quasiball. Seittenranta [20] proved that the inequality $j_D \leq \alpha_D$ holds if D is convex and Hästö [13] gave a geometric characterization for domains in which the inequality $\frac{1}{K}j_D \leq \alpha_D$ holds (for a definition of the j_D metric see, e.g., [10, 11, 13]).

Another aspect of the Apollonian metric, which has recently been studied, is its conformality [16]. We say that the Apollonian metric of a domain $D \subset \overline{\mathbb{R}^n}$ is conformal at a point $x \in D$ if

$$\lim_{r \rightarrow 0} \frac{\max\{\alpha(x, y) : y \in S^{n-1}(x, r)\}}{\min\{\alpha(x, y) : y \in S^{n-1}(x, r)\}} = 1.$$

It turns out that the notion of conformality of the Apollonian metric is intimately related to the notion of a constant width set, an object which has been studied by geometers for several centuries (see e.g. [8] and the references therein). Namely, the Apollonian metric α_D of a domain $D \subset \overline{\mathbb{R}^n}$ is conformal either at every point, at exactly one point or at no point of D [16, Theorem 2]. The first case occurs if and only if D is a ball. For the second case we have the following result [16, Theorem 3]: the Apollonian metric α_D is conformal at one point if and only if, up to a Möbius transformation, D is the complement of a convex body of constant width. The relation between the convex bodies of constant width and the Apollonian metric is being investigated more closely by the second author and will be reported elsewhere [5].

2. NOTATION AND CONVENTIONS

We denote by \mathbb{R}^n the n -dimensional euclidean space and by $\{e_1, e_2, \dots, e_n\}$ its standard basis. Open balls and spheres of radius $r > 0$ centered at $x \in \mathbb{R}^n$ are denoted by $B^n(x, r)$ and $S^{n-1}(x, r)$, respectively. The closed segment between two

points x and y in \mathbb{R}^n is denoted by $[x, y]$. For $x \in \mathbb{R}^n \setminus \{0\}$ we set

$$[x, \infty] = \{tx : t \geq 1\} \cup \{\infty\}.$$

The diameter of a set A is denoted by $\text{diam}(A)$ and its cardinality by $\text{card}(A)$. By D we always denote an open set in $\overline{\mathbb{R}^n}$ with at least two boundary points.

The Möbius space $\overline{\mathbb{R}^n} = \mathbb{R}^n \cup \{\infty\}$ is the one-point compactification of \mathbb{R}^n equipped with the chordal metric

$$d(x, y) = \begin{cases} \frac{|x-y|}{\sqrt{1+|x|^2}\sqrt{1+|y|^2}} & \text{if } x, y \in \mathbb{R}^n, \\ \frac{1}{\sqrt{1+|x|^2}} & \text{if } y = \infty. \end{cases}$$

For a set A in \mathbb{R}^n or $\overline{\mathbb{R}^n}$ the topological operations \overline{A} (closure) and ∂A (boundary) are always taken with respect to $\overline{\mathbb{R}^n}$. The *cross-ratio* of a quadruple a, b, c, d of points in $\overline{\mathbb{R}^n}$ with $a \neq b$ and $c \neq d$ is defined by

$$|a, b, c, d| = \frac{d(a, c)d(b, d)}{d(a, b)d(c, d)} = \frac{|a - c||b - d|}{|a - b||c - d|},$$

where the latter equality holds with the understanding that that $|a - \infty|/|b - \infty| = 1$ for all $a, b \in \mathbb{R}^n$. A homeomorphism $f: \overline{\mathbb{R}^n} \rightarrow \overline{\mathbb{R}^n}$ is a *Möbius transformation* if

$$|f(a), f(b), f(c), f(d)| = |a, b, c, d|$$

for all quadruples of distinct points a, b, c, d in $\overline{\mathbb{R}^n}$. Properties of Möbius mappings will play a central role in most of our arguments, and are too numerous to be reviewed here. The reader is referred to an introduction on the geometry of Möbius mappings (e.g. [2]) for further reference.

By a ball in $\overline{\mathbb{R}^n}$ we mean the image of $B^n(0, 1)$ under a Möbius transformation of $\overline{\mathbb{R}^n}$, i.e. a euclidean ball, the complement of a closed euclidean ball or a half-space. Two distinct points $x, y \in \overline{\mathbb{R}^n}$ are called *inversive* with respect to the sphere S in $\overline{\mathbb{R}^n}$ if $y = h(x)$ where h is the inversion in S .

3. THE APOLLONIAN METRIC

For distinct $x, y \in \overline{\mathbb{R}^n}$ and $k \in (0, +\infty)$ the *Apollonian sphere* $S_\alpha(x, y; k)$ and the *Apollonian ball* $B_\alpha(x, y; k)$ of radius k about x with respect to y are defined as

$$\left\{ w \in \overline{\mathbb{R}^n} : \frac{d(w, x)}{d(w, y)} = k \right\} \quad \text{and} \quad \left\{ w \in \overline{\mathbb{R}^n} : \frac{d(w, x)}{d(w, y)} < k \right\},$$

respectively. Observe that the Apollonian spheres are spheres in $\overline{\mathbb{R}^n}$. The points x and y are inversive with respect to $S_\alpha(x, y; k) = S_\alpha(y, x; 1/k)$. Given an open set $D \subset \overline{\mathbb{R}^n}$ with $\text{card}(\partial D) \geq 2$, for each pair of distinct points $x, y \in D$ we let k_{xy} be the radius of the largest Apollonian ball in D about x with respect to y . The Apollonian ball $B_\alpha(x, y; k_{xy})$ is called the *maximal Apollonian ball* in D about x with respect to y and is denoted by $B_\alpha(x, y)$, similarly $S_\alpha(x, y) = S_\alpha(x, y; k_{xy})$. We easily see that

$$k_{xy} = \min_{w \in \partial D} \frac{d(w, x)}{d(w, y)} \quad \text{and} \quad k_{xy}k_{yx} \leq 1$$

for all $x, y \in D$. The *Apollonian distance* $\alpha_D(x, y)$ between the points x and y is defined as

$$\alpha_D(x, y) = \max_{w, z \in \partial D} \log \left(\frac{|y - z||x - w|}{|x - z||y - w|} \right) = \log \left(\frac{1}{k_{xy}} \cdot \frac{1}{k_{yx}} \right).$$

This distance is a metric if and only if $\overline{\mathbb{R}^n} \setminus D$ is not contained in a sphere in $\overline{\mathbb{R}^n}$ [2, Theorem 1.1]. For each pair of distinct points $x, y \in D$ we define an *extremal subset* of ∂D about x with respect to y as $E(x, y) = \partial D \cap S_\alpha(x, y)$. Notice the implicit dependence of $E(x, y)$ on ∂D ; this will cause no confusion, since D will always be clear from context. The points in the extremal subset are called *extremal points*.

We say that a homeomorphism $f: D \rightarrow f(D) \subset \overline{\mathbb{R}^n}$ is an *Apollonian isometry* if

$$\alpha_D(x, y) = \alpha_{f(D)}(f(x), f(y))$$

for all $x, y \in D$. It follows directly from the definitions that every Möbius mapping is an Apollonian isometry. The purpose of this paper is to show that in many cases these are all the isometries.

Let us indicate why we study only homeomorphisms and domains with sufficiently large boundary. The reason is that if the boundary is too small, then the Apollonian distance is not a metric and strange isometries exist. One easily sees that the examples used by Hästö and Lindén [14, Example 4] to show that a related metric, the so-called half-apollonian metric, has irregular isometries work for the Apollonian metric as well. More precisely, if the boundary of D is properly contained in a sphere, then there exist (euclidean) discontinuous Apollonian isometries of α_D . If the boundary of D is contained in an $(n - 2)$ -dimensional sphere, then there exist continuous Apollonian isometries which are not Möbius mappings.

The next result shows that looking only at homeomorphisms is no restriction in the case where the complement is not contained in a sphere.

Proposition 3.1. *Suppose that $f: D \rightarrow \overline{\mathbb{R}^n}$ is such that*

$$\alpha_D(x, y) = \alpha_{f(D)}(f(x), f(y))$$

for all $x, y \in D$ and that $\overline{\mathbb{R}^n} \setminus D$ is not contained in a sphere. Then f is an embedding.

Proof. Since α_D is a metric we see that f is injective and $\alpha_{f(D)}$ is also a metric. Fix $x \in D$ and let $x_i \in D$ be a sequence of points tending to x . Then $\alpha_D(x, x_i) \rightarrow 0$. Suppose that $f(x_i) \not\rightarrow f(x)$. Since $\overline{f(D)}$ is compact there exists a subsequence (x_{i_k}) such that $f(x_{i_k})$ converges to a point z in $\overline{f(D)}$. If $z \in \partial f(D)$, then $\alpha_{f(D)}(f(x_{i_k}), f(x)) \rightarrow \infty$, a contradiction. Otherwise

$$\alpha_{f(D)}(f(x_{i_k}), f(x)) \rightarrow \alpha_{f(D)}(z, f(x)).$$

Since $\alpha_{f(D)}$ is metric, $\alpha_{f(D)}(z, f(x)) > 0$, again a contradiction. Therefore f is continuous. By symmetry, f^{-1} is continuous as well. \square

4. GEODESICS

Since we are only interested in Apollonian geodesics, we will call an arc $\gamma \subset D$ a *geodesic* if

$$\alpha_D(x, y) = \alpha_D(x, z) + \alpha_D(z, y)$$

for each suitably ordered triple of points $x, z, y \in \gamma$. (It is also possible to treat geodesics in terms of a directed density related to the Apollonian metric, see [12].) Let $\{a, b\}$ be the endpoints of γ . We say that the geodesic γ is a *geodesic line, ray* or *segment* if two, one or none of a, b are in ∂D . There is a geodesic connecting two arbitrary points of D if and only if D is a ball.

The following lemma will be used throughout this section. A similar result was announced already by Barbilian [1, Satz I] and was proven by Gehring and Hag

for the planar case, see [9, Lemma 3.18]. Since their proof carries over to higher dimensions, it is not included here.

Lemma 4.1. *Let x, y, z be distinct points in D . Then*

$$\alpha_D(x, y) = \alpha_D(x, z) + \alpha_D(z, y)$$

if and only if $E(x, y) \subset E(x, z) \cap E(z, y)$ and $E(y, x) \subset E(y, z) \cap E(z, x)$.

The main result of this section is Theorem 4.6, which shows that through every point of an open set D there passes a pseudogeodesic line and, in particular, a geodesic segment. We begin with the following preliminary result.

Lemma 4.2. *Let γ be a curve in D with end-point $a \in \partial D$. Let $x, y \in \gamma$ be such that for all $z \in \gamma$ in the order a, z, x, y we have*

$$\alpha_D(z, y) = \alpha_D(z, x) + \alpha_D(x, y).$$

Then $a \in E(x, y)$.

Proof. Suppose that $a \notin E(x, y)$. By compactness of $E(x, y)$, we have

$$r = \frac{1}{2} \min\{d(w, a) : w \in E(x, y)\} > 0.$$

Then $d(z, w) > r$ for all $z \in \gamma$ with $d(z, a) < r$ and all $w \in E(x, y)$. Now choose a point $z \in \gamma$ such that $d(z, a) < r$ and $rd(y, a) > d(z, a)$. Since $d(y, w) \leq 1$ by definition, we find for every $w \in E(x, y)$ that

$$\frac{d(y, a)}{d(z, a)} > \frac{1}{r} > \frac{1}{d(z, w)} > \frac{d(y, w)}{d(z, w)},$$

i.e. $w \notin E(z, y)$. Hence $E(z, y) \cap E(x, y) = \emptyset$. But as $\alpha_D(z, y) = \alpha_D(z, x) + \alpha_D(x, y)$, Lemma 4.1 implies that $E(z, y) \subset E(x, y) \cap E(z, x)$, a contradiction. Thus, $a \in E(x, y)$. \square

The proof of the previous lemma was very detailed, but not so intuitive or geometric. As we go on, the proofs will become increasingly geometric, we will rely less on explicitly writing down the sets $E(x, y)$ and think more in terms of intersections of maximal Apollonian spheres.

As an immediate corollary of Lemma 4.2 we obtain

Corollary 4.3. *Let γ be a geodesic line in D with distinct endpoints $w_1, w_2 \in \partial D$. Then $w_1 \in E(x, y)$ and $w_2 \in E(y, x)$ for all $x, y \in \gamma$ such that x lies between w_1 and y in $\overline{\gamma}$.*

The following lemma gives us an existence criterion for a geodesic segment passing through a given point.

Lemma 4.4. *Let x, z, y be distinct points in D with*

$$\alpha_D(x, y) = \alpha_D(x, z) + \alpha_D(z, y)$$

and let γ be the circular arc joining x and y and containing z . Then

- (1) *there is a closed non-degenerate subarc S of γ containing z in its interior which is a geodesic;*
- (2) *for all $x', y' \in S$ such that the points x, x', y', y lie on γ in this order, we have $E(x, y) \subset E(x', y')$ and $E(y, x) \subset E(y', x')$;*

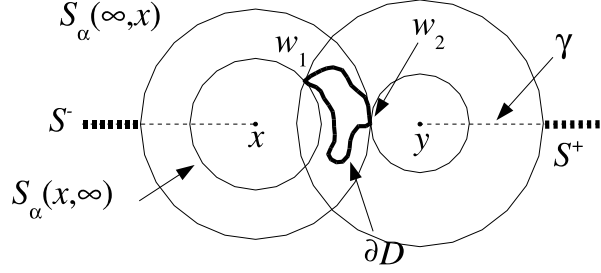


Figure 1.

- (3) for each x' in $\gamma \setminus \{x\}$ lying between x and z , there exists a closed non-degenerate subarc S' of S having z as one of its endpoints such that

$$\alpha_D(x', z) = \alpha_D(x', z') + \alpha_D(z', z) \quad \text{for all } z' \in S'.$$

Proof. Without loss of generality we assume that $x = -e_1$, $y = e_1$ and $z = \infty$, see Figure 1. Then $\gamma = [-e_1, \infty] \cup [e_1, \infty]$. Let $w_1 \in E(x, y)$ and $w_2 \in E(y, x)$ be arbitrary points. Then $w_1 \in E(-e_1, \infty) \cap E(\infty, e_1)$ and $w_2 \in E(\infty, -e_1) \cap E(e_1, \infty)$ by Lemma 4.1. Notice that the Apollonian ball $B_\alpha(w, \infty)$ is just a euclidean ball centered at w and $B_\alpha(\infty, w)$ is the complement of a closed euclidean ball centered at w . Hence we have $\partial D \subset \overline{B^n}(-e_1, |w_2 + e_1|) \setminus B^n(-e_1, |w_1 + e_1|)$ and similarly, $\partial D \subset \overline{B^n}(e_1, |w_1 - e_1|) \setminus B^n(e_1, |w_2 - e_1|)$.

We will show that $S = S^- \cup S^+$ is the required geodesic segment, where

$$S^- = [-(1 + |w_2 + e_1|)e_1, \infty] \quad \text{and} \quad S^+ = [(1 + |w_1 - e_1|)e_1, \infty].$$

Let $re_1, se_1, te_1 \in S$ with $r < s < t$. Due to symmetry we can assume that $re_1, se_1 \in S^-$. We also suppose that $te_1 \in S^+$ (the other case is similar). Then we conclude that

$$w_1 \in E(se_1, re_1) \cap E(se_1, te_1) \cap E(re_1, te_1);$$

for instance, to see that $w_1 \in E(se_1, re_1)$, we note that $S_\alpha(se_1, re_1)$ is a sphere centered the negative e_1 -axis, which crosses this axis between re_1 and se_1 . It is easy to see such a sphere intersects ∂D in w_1 . The other inclusions are proved similarly, and so is:

$$w_2 \in E(se_1, re_1) \cap E(te_1, se_1) \cap E(te_1, re_1).$$

This proves part (ii) and, in particular, implies that

$$\begin{aligned} \alpha_D(se_1, te_1) &= \log \frac{|se_1 - w_2||te_1 - w_1|}{|se_1 - w_1||te_1 - w_2|} \\ &= \log \left(\frac{|se_1 - w_2||re_1 - w_1|}{|se_1 - w_1||re_1 - w_2|} \frac{|re_1 - w_2||te_1 - w_1|}{|re_1 - w_1||te_1 - w_2|} \right) \\ &= \alpha_D(se_1, re_1) + \alpha_D(re_1, te_1). \end{aligned}$$

Hence S is a geodesic, completing the proof of part (i).

Finally, to prove part (iii), let x' be as in the lemma. If x' is on the geodesic segment S , then the claim is obvious. Otherwise let x^* be the inversion of x' in $S^{n-1}(-e_1, |w_2 + e_1|)$. Then similar observations as above imply that the subarc S' of S^- joining x^* to z is the required arc. \square

We are now ready to introduce our main tool, the pseudogeodesic line:

Definition 4.5. Let $\gamma \subset D$ be a curve joining two boundary points a and b . Suppose that $z \in \gamma$ and that there exists a non-degenerate geodesic subcurve S of γ containing z in its interior, with the additional property that for every $x \in \gamma$ there exists a closed non-degenerate subcurve S' of S with one endpoint at z such that $a_D(x, z) = a_D(x, y) + a_D(y, z)$ for all $y \in S'$. Then γ is called a *pseudogeodesic line through z* .

Observe that pseudogeodesic lines include geodesic lines and that pseudogeodesic lines are mapped to pseudogeodesic lines by Apollonian isometries. The next theorem shows why pseudogeodesic lines are so useful to us: through every point there is at least one.

Theorem 4.6. *Let $D \subset \overline{\mathbb{R}^n}$ be open with $\text{card}(\partial D) \geq 2$ and let $x \in D$. Then*

- (1) *there exists a pseudogeodesic line through x ;*
- (2) *there exists a ball $B_x \subset D$ containing x and a geodesic line.*

Proof. Without loss of generality we can assume that $x = \infty$, $\text{diam}(\partial D) = 2$ and that $-e_1, e_1 \in \partial D$ are diametrical points of ∂D . Then the complement of $B^n(-e_1, 2) \cap B^n(e_1, 2)$ is contained in D and one can easily observe that

$$\alpha_D(-se_1, te_1) = \alpha_D(-se_1, \infty) + \alpha_D(\infty, te_1)$$

for all $s, t \in (1, +\infty)$. The proof of (1) is then completed using Lemma 4.4.

Next let B be the unique ball of smallest radius containing the complement of D (see, for instance, [4, Theorem 11.5.8, p. 357]). Then one can easily observe that $\text{card}(\partial B \cap \partial D) \geq 2$. Define $B_x = \overline{\mathbb{R}^n} \setminus B$. Let now $a, b \in \partial B_x \cap \partial D$ be two distinct points and let γ be a circular arc in B_x orthogonal to ∂B_x at a and b . Then a and b are the extremal points for every pair of points on γ , and so it is easy to see that γ is a geodesic. Thus B_x is the ball sought for in (2). \square

Using this results we can easily prove a result which Barbilian stated in the special case of a domain with regular boundary [1, Satz III].

Corollary 4.7. *Let $D \subset \overline{\mathbb{R}^n}$ be an open set with $\text{card}(\partial D) \geq 2$. Then there exist infinitely many geodesic lines in D .*

Proof. Using the second part of the previous theorem we find a ball B_x containing a geodesic. If possible, we choose a point y from the set $D \setminus \overline{B_x}$ and repeat the above argument to produce another geodesic line and continue this process. If the process stops after a finite number of steps, then D is the union of finitely many balls and there are infinitely many geodesic lines in at least one of these balls. Otherwise, this process will generate infinitely many geodesic lines. \square

5. ISOMETRIES NEAR THE BOUNDARY

We define $\delta(x) = d(x, \partial D)$ and $\delta'(y) = d(y, \partial f(D))$, where $x \in D \subset \mathbb{R}^n$, f is a function defined on D and $y \in f(D)$. We start by looking at a kind of dilatation on the boundary.

Lemma 5.1. *Let $D \subset \overline{\mathbb{R}^n}$ be an open set with at least two boundary points and let $f: D \rightarrow \overline{\mathbb{R}^n}$ be an Apollonian isometry. Assume that ∂D and $\partial f(D)$ are bounded.*

If $z \in \partial D$ and (z_i) is a sequence of points in D tending to z , then

$$h_f(z) = \lim_{i \rightarrow \infty} \frac{\delta'(f(z_i))}{\delta(z_i)}$$

exists and depends only on the point z , not on the sequence (z_i) . The mapping f extends continuously to a map of \overline{D} and

$$h_f(z)h_f(w) = \frac{|f(z) - f(w)|^2}{|z - w|^2}$$

for all $z, w \in \partial D$.

Proof. Since we will only be interested in points near the boundary, we assume that all points in this proof are finite. Let z and (z_i) be as in the statement of the lemma and let w be a second boundary point with a sequence (w_i) of points in D approaching w . Let us choose from (z_i) a subsequence such that $f(z_i)$ converges to a boundary point fz (in $\overline{\mathbb{R}^n}$) and from (w_i) a subsequence such that $f(w_i)$ converges to a boundary point fw , and denote these subsequences again by (z_i) and (w_i) , respectively.

Using the triangle inequalities $|x - y| - |x - a| \leq |y - a| \leq |x - y| + |x - a|$ we find that

$$\frac{|x - y|}{\delta(x)} - 1 \leq \max_{a \in \partial D} \frac{|y - a|}{|x - a|} \leq \frac{|x - y|}{\delta(x)} + 1.$$

The same inequality holds for x and y interchanged. Hence we get

$$(5.2) \quad \begin{aligned} \log \left(\frac{|x - y|}{\delta(x)} - 1 \right) \left(\frac{|x - y|}{\delta(y)} - 1 \right) &\leq \alpha_D(x, y) \\ &\leq \log \left(\frac{|x - y|}{\delta(x)} + 1 \right) \left(\frac{|x - y|}{\delta(y)} + 1 \right). \end{aligned}$$

Fix $x \in D$ far enough from w that $2 \leq |x - w_i|/\delta(x)$ for all sufficiently large i . By taking i larger still, we also assume that $|x - w| \leq 2(|x - w_i| - \delta(w_i))$ and $|f(x) - f(w_i)| + \delta(w_i) \leq 2|f(x) - fw|$. Then we have

$$\begin{aligned} &\log \left(\frac{1}{2}|x - w| \right) - \log \delta(w_i) \\ &\leq \log (|x - w_i| - \delta(w_i)) - \log \delta(w_i) + \log \left(\frac{|x - w_i|}{\delta(x)} - 1 \right) \\ &\leq \alpha_D(x, w_i) = \alpha_{f(D)}(f(x), f(w_i)) \\ &\leq \log (|f(x) - f(w_i)| + \delta'(f(w_i))) - \log \delta'(f(w_i)) + \log \left(\frac{|f(x) - f(w_i)|}{\delta'(f(x))} + 1 \right) \\ &\leq \log (2|f(x) - fw|) - \log \delta'(f(w_i)) + \log \left(\frac{2|f(x) - fw|}{\delta'(f(x))} + 1 \right). \end{aligned}$$

Therefore the sequence $\delta'(f(w_i))/\delta(w_i)$ is bounded. Similarly we conclude that this sequence is bounded away from 0. Thus we can choose a subsequence of (w_i) , denoted again by (w_i) , such that

$$\delta'(f(w_i))/\delta(w_i) \rightarrow C_w \in (0, \infty).$$

Using again the fact that f is an Apollonian isometry and (5.2) we get

$$\begin{aligned} \log \left(\frac{|f(z_i) - f(w_i)|}{\delta'(f(z_i))} + 1 \right) \left(\frac{|f(z_i) - f(w_i)|}{\delta'(f(w_i))} + 1 \right) &\geq \alpha_{f(D)}(f(z_i), f(w_i)) \\ &= \alpha_D(z_i, w_i) \geq \log \left(\frac{|z_i - w_i|}{\delta(z_i)} - 1 \right) \left(\frac{|z_i - w_i|}{\delta(w_i)} - 1 \right). \end{aligned}$$

We take the exponential function of both sides of this inequality and rearrange (all terms are positive for large enough i) to get

$$\frac{\delta(w_i)}{\delta'(f(w_i))} \frac{|f(z_i) - f(w_i)| + \delta'(f(z_i))}{|z_i - w_i| - \delta(z_i)} \frac{|f(z_i) - f(w_i)| + \delta'(f(w_i))}{|z_i - w_i| - \delta(w_i)} \geq \frac{\delta'(f(z_i))}{\delta(z_i)}$$

In this inequality we let $z_i \rightarrow z$ and $w_i \rightarrow w$, and note that $|z_i - w_i|$ is bounded away from zero, whereas all the δ and δ' functions tend to zero. This gives us the inequality

$$\limsup_{i \rightarrow \infty} \frac{\delta'(f(z_i))}{\delta(z_i)} \leq \frac{1}{C_w} \frac{|fz - fw|^2}{|z - w|^2}.$$

If $|fz - fw| = 0$, then the following inequality is trivial, otherwise, a similar argument as before works, so in either case we have

$$\liminf_{i \rightarrow \infty} \frac{\delta'(f(z_i))}{\delta(z_i)} \geq \frac{1}{C_w} \frac{|fz - fw|^2}{|z - w|^2}.$$

Combining these inequalities gives

$$(5.3) \quad \lim_{i \rightarrow \infty} \frac{\delta'(f(z_i))}{\delta(z_i)} = \frac{1}{C_w} \frac{|fz - fw|^2}{|z - w|^2}.$$

Since the left hand side does not depend on w , we conclude that the right hand side must also be independent of w . Therefore we have shown that the limit can depend only on z and fz .

Suppose now that $f(z_i)$ had more than one accumulation point, say fz^1 and fz^2 , and that (z_i^j) are subsequences of (z_i) such that $\lim_i f(z_i^j) = fz^j$ ($j = 1, 2$). Then there exists a constant $K > 0$ such that

$$\begin{aligned} \exp(\alpha_{f(D)}(f(z_i^1), f(z_i^2))) &\geq \left(\frac{|f(z_i^1) - f(z_i^2)|}{\delta'(f(z_i^1))} - 1 \right) \left(\frac{|f(z_i^1) - f(z_i^2)|}{\delta'(f(z_i^2))} - 1 \right) \\ &\geq K \{\delta'(f(z_i^1))\delta'(f(z_i^2))\}^{-1} \end{aligned}$$

for large enough i , since $|f(z_i^1) - f(z_i^2)| \rightarrow |fz^1 - fz^2| > 0$. On the other hand

$$\exp(\alpha_D(z_i^1, z_i^2)) \leq \left(\frac{|z_i^1 - z_i^2|}{\delta(z_i^1)} + 1 \right) \left(\frac{|z_i^1 - z_i^2|}{\delta(z_i^2)} + 1 \right) = o(\{\delta(z_i^1)\delta(z_i^2)\}^{-1}),$$

because $|z_i^1 - z_i^2| \rightarrow 0$. But we have already shown that $\delta'(f(z_i^j))/\delta(z_i^j) \rightarrow C_j \in (0, \infty)$, so these inequalities contradict the fact that f is an Apollonian isometry.

Therefore fz is unique, and so the limit in (5.3) depends only on z , as claimed, and equals $h_f(z)$. Of course, this argument also applies to w , and so we see that $C_w = h_f(w)$. We extend f to ∂D by defining $f(z) = fz$ for boundary points z . Then (5.3) is just the equation in the statement of the lemma.

It remains to check that the extension is continuous. If $z_i \in D$ and $z_i \rightarrow z \in \partial D$ then it is clear by construction that $f(z_i) \rightarrow f(z)$. So it remains to consider $z_i \in \partial D$ with $z_i \rightarrow z \in \partial D$. But like when we concluded that C_w is bounded, we find that h_f

has a uniform upper bound, so f is Lipschitz continuous on the boundary. Therefore f is continuous on \overline{D} . \square

Recall the following fact about inversions (e.g. [2, p. 26]). Recall also that every inversion is its own inverse.

Lemma 5.4. *Let g be an inversion in the sphere $S^{n-1}(z, r)$. Then*

$$|g(x) - g(y)| = \frac{r^2|x - y|}{|x - z||y - z|}.$$

The inversion is conformal and the dilatation is given by $\frac{r^2}{|x - z|^2}$.

Using inversions suitably, we conclude that Lemma 5.1 holds without the restriction on the boundary:

Lemma 5.5. *Let $D \subset \overline{\mathbb{R}^n}$ be an open set with at least two boundary points and let $f: D \rightarrow \overline{\mathbb{R}^n}$ be an Apollonian isometry. If $z \in \partial D$ and (z_i) is a sequence of points in D tending to z , then*

$$h_f(z) = \lim_{i \rightarrow \infty} \frac{\delta'(f(z_i))}{\delta(z_i)}$$

exists and depends only on the point z , not on the sequence (z_i) . The mapping f extends to a continuous map of \overline{D} and

$$h_f(z)h_f(w) = \frac{|fz - fw|^2}{|z - w|^2}$$

for all finite boundary points z, w .

Proof. Suppose that $x \in D$ and $y \in f(D)$ and let g_1 and g_2 be inversions in the spheres $S(x, 1)$ and $S(y, 1)$. Then the set $D' = g_1(D)$ and the mapping $f' = g_2 \circ f \circ g_1$ satisfy the assumptions of Lemma 5.1. Therefore f' extends to a continuous mapping of $\overline{D'}$ and

$$(5.6) \quad h_{f'}(z)h_{f'}(w) = \frac{|f'z - f'w|^2}{|z - w|^2}$$

for $z, w \in \partial D'$. But Möbius mappings are continuous on $\overline{\mathbb{R}^n}$, so this implies that f extends continuously to the boundary of D . One checks easily using Lemma 5.4 that formula (5.6) is Möbius invariant, so it holds for f , too. \square

Theorem 5.7. *Let $D \subset \overline{\mathbb{R}^n}$ be an open set with at least $n+1$ boundary points, which are not contained in an hyperplane. Let $f: D \rightarrow \overline{\mathbb{R}^n}$ be an Apollonian isometry and extend it to \overline{D} continuously. Then, up to composition by Möbius maps, $f(D) = D$ and $f|_{\partial D} = id_{\partial D}$.*

Proof. Consider the case $n = 2$ first. Let z_1, z_2 and z_3 be three distinct boundary points of D . By composing f with Möbius mappings, we may assume that $f(z_i) = z_i$ for these points, and that the points do not lie on a line. By Lemma 5.5 this implies that

$$h_f(z_1)h_f(z_2) = h_f(z_1)h_f(z_3) = h_f(z_2)h_f(z_3) = 1,$$

from which we conclude that $h_f(z_1) = h_f(z_2) = h_f(z_3) = 1$. Then, by Lemma 5.5,

$$(5.8) \quad |f(a) - z_i| = \sqrt{h_f(a)}|a - z_i|$$

for $i = 1, 2, 3$ and every $a \in \partial D$.

Let $a \in \partial D$. If $h_f(a) = 1$, then the distance of a to the three points z_1, z_2 and z_3 is preserved, so $f(a) = a$. Suppose next that $h_f(a) \neq 1$. Let us write $K = \sqrt{h_f(a)}$ and define

$$C = \left\{ x \in \overline{\mathbb{R}^n} : \frac{|f(a) - x|}{|a - x|} = K \right\}.$$

Now C is a circle, which by (5.8) contains z_1, z_2 and z_3 . Moreover, a and $f(a)$ are inverse points with respect to this circle. But z_1, z_2 and z_3 define a unique circle, so we have shown that for every $a \in \partial D$ either $f(a) = a$ or $f(a) = a^*$ (the inverse of a in C). We still have to show that if f acts as an inversion on one boundary point, then it acts as an inversion on all boundary points. So suppose $f(a) = a^* \neq a$ and $f(b) = b, a, b \in \partial D$. Since b is fixed, we easily see that $h_f(b) = 1$. Hence $\frac{|f(a) - b|}{|a - b|} = K$, so that b lies on C . Thus f acts as an inversion on all boundary points in this case. Otherwise, f is just the identity for all boundary points, and, in either case, the identity up to a Möbius map.

In the higher dimensional case we will show that we can assume that the $n + 1$ boundary points are fixed: as before we start by fixing three boundary points. Assume then inductively that we have fixed $k \in [3, n]$ boundary points contained in the $(k - 1)$ -dimensional plane H , but not in any $(k - 2)$ -dimensional plane. Let $S = S^{k-2}(x, r)$ be the $(k - 2)$ -sphere in H containing the points and let u be a unit vector orthogonal to H . Then the sphere

$$S_t = S^{n-1}(x + tu, \sqrt{r^2 + t^2})$$

contains S , so the inversion in S_t fixes our k points for every t . Pick one boundary point a not in H . We easily see using Lemma 5.4 that we can choose t so that modifying f by composition with the inversion in S_t makes $h_f(a) = 1$. Then f acts as an isometry on the set formed by our k points and a , so we can compose f with rotations and reflections which leave H pointwise fixed so that $f(a) = a$. Now we just add this point to our collection and have thus completed the inductive step.

Once we have found $n + 1$ boundary points not all lying in the same hyperplane we can argue as in the second paragraph of this proof to show that f is either the identity or an inversion in the sphere containing our boundary points. \square

Using the theorem we can directly prove the following result, which has previously been proven by Gehring and Hag [9]:

Corollary 5.9. *If $f: B \rightarrow \mathbb{R}^n$ is an Apollonian isometry from a ball, then f is a Möbius mapping.*

Proof. By Theorem 5.7 we know that $f(B)$ is a ball. Since the Apollonian metric equals the hyperbolic metric on the ball, the claim follows. \square

6. THE MAIN RESULT

Let us start by formally defining what we mean by a regular boundary point in the context of the Apollonian metric.

Definition 6.1. Let $D \subset \mathbb{R}^n$ be an open set. A boundary point $w \in \partial D$ is irregular if there exist two distinct balls in D so that the intersection of their boundaries contains w and some other point. Boundary points which are not irregular are *regular*. If every boundary point of the open set D is regular, then we say that D has *regular boundary*.

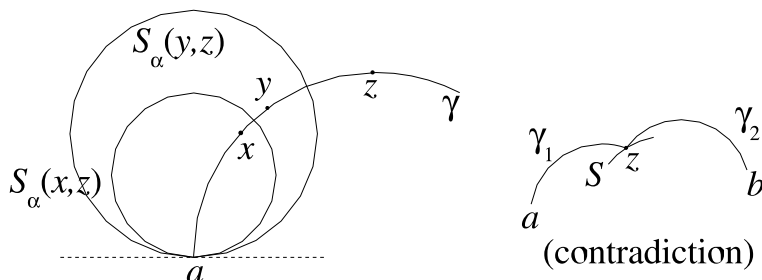


Figure 2.

For example a domain with C^1 boundary or a convex domain has regular boundary. Notice that being a regular boundary point is a Möbius invariant feature. This follows, since it is easily seen that being irregular is Möbius invariant. We record the following immediate fact.

Lemma 6.2. *If w is a regular boundary point, then there exists a unique line U_w through w so that every ball in D whose boundary contains w is centered on U_w .*

Lemma 6.3. *Let $D \subset \overline{\mathbb{R}^n}$ be an open set with regular boundary. Then every pseudogeodesic line is just a circular arc and there are at most two pseudogeodesic lines joining any pair of boundary points.*

Proof. Let γ be a pseudogeodesic line through z and joining $a, b \in \partial D$. Choose $x, y \in \gamma$ such that $\alpha_D(x, z) = \alpha_D(x, y) + \alpha_D(y, z)$. The maximal Apollonian spheres $S_\alpha(x, y)$ and $S_\alpha(y, z)$ meet at some boundary point p . But since p is a regular boundary point, this means that it is their only intersection, see Figure 2. Therefore $E(x, z) = \{p\}$, by Lemma 4.1. Moreover, for all y' closer to z than y , we easily see that we still have $S_\alpha(y', z) \cap \partial D = \{p\}$. But then we can take x' close to a and y' close to z with $\alpha_D(x', z) = \alpha_D(x', y') + \alpha_D(y', z)$, and argue as in Lemma 4.2 to show that $p = a$. Thus we have shown that $E(x, y) = S_\alpha(x, z) \cap \partial D = \{a\}$ for every $x \in \gamma$ between a and z .

Recall that x and z are inverse with respect to the maximal Apollonian sphere. Consider the locus of points which are the inversion of z in a sphere through a with center on U_a (which is as in Lemma 6.2). It is well-known that such a locus is the arc of a circle which is tangent to U_a at a and passes through z . So this arc of the circle is the pseudogeodesic on the a -side, call it γ_1 .

By a similar argument we conclude that the pseudogeodesic on the b -side is the arc of a circle, call it γ_2 . Finally, we can apply the same argument to the geodesic segment S on the pseudogeodesic so it too is an arc of a circle. As can be seen in the right part of Figure 2, this means that γ_1 and γ_2 must be part of the same arc of a circle. Therefore the pseudogeodesic lies on the unique circle C through a and b , tangent to U_a and U_b at these points, so there are at most two pseudogeodesic connecting a and b , namely the two components of $C \setminus \{a, b\}$. \square

Theorem 6.4. *Let $D \subset \overline{\mathbb{R}^n}$ be an open set with regular boundary. If $f: D \rightarrow \overline{\mathbb{R}^n}$ is an Apollonian isometry, then f is the restriction to D of a Möbius mapping.*

Proof. If ∂D is a sphere the the conclusion follows easily from Corollary 5.9. Otherwise ∂D is not contained in a sphere in $\overline{\mathbb{R}^n}$, for if this were the case, then ∂D

would contain irregular points. Using Theorem 5.7 we may assume that $f(D) = D$, that f extends continuously to the boundary and that it fixes every point of the boundary.

Fix $z \in D$. By Theorem 4.6 and Lemma 6.3 there exists a circular pseudogeodesic line γ through z . Since the pseudogeodesic lines are defined only in terms of the Apollonian metric it is clear that $f(\gamma)$ is also a pseudogeodesic line. Since $f(\gamma)$ has the same end-points as γ it follows from Lemma 6.3 that $f(\gamma) = \gamma$ or $f(\gamma)$ is the complement of γ in the circle through γ . Let us assume first that $f(\gamma) = \gamma$.

Consider points w of γ tending to a . Since a and b are the extremal points for z and w we get

$$\alpha_D(z, w) + \log \delta(w) = \log \left(|w - b| \frac{|z - a|}{|z - b|} \right) + \log \frac{\delta(w)}{|w - a|}.$$

We do a similar computation for $\alpha_D(f(z), f(w))$, using the same two extremal points. Now $\delta(w)/\delta(f(w)) \rightarrow 1$ since $h_f(w) = 1$, and $|w - a|/\delta(w) \rightarrow 1$ and $|f(w) - a|/\delta(f(w)) \rightarrow 1$ since the pseudogeodesic is the arc of a circle orthogonal to the boundary. Thus we find that

$$1 = \lim_{w \rightarrow a} \frac{\delta(f(w)) \exp \alpha_D(f(z), f(w))}{\delta(w) \exp \alpha_D(z, w)} = \frac{|a - b| \frac{|f(z) - a|}{|f(z) - b|}}{|a - b| \frac{|z - a|}{|z - b|}} = \frac{|f(z) - b| |z - a|}{|f(z) - a| |z - b|}.$$

Since we know that $f(z)$ must lie on the circular arc $\gamma = f(\gamma)$, this equality implies that z is fixed by f . Therefore we are done if $f(\gamma) = \gamma$ for all pseudogeodesics.

Let us partition D into those points which are fixed by f , F , and those which are not, N . Since f is continuous it is clear that F is closed (in D). We will show that N is closed, too. This is clear if N has no accumulation points. If it has such points, let $z \in D$ be one of them. Then we chose a sequence (z_i) of points from $N \cap B^n(z, \delta(z)/2)$ which tend to z . Every z_i is on a pseudogeodesic which gets inverted under f (otherwise z_i would be fixed, by the previous argument). Since $\delta(z_i) \geq \delta(z)/2$ we easily see that $|f(z_i) - z_i| > \delta(z)/2$. Since f is continuous this means that $|f(z) - z| \geq \delta(z)/2$. Therefore $f(z) \neq z$, so $z \in N$ and N is closed.

We have thus partitioned D in two closed sets, which means that every component of D must lie completely in one of these sets. By Theorem 4.6(ii) there exists a geodesic line γ in every component of D . Suppose that γ is mapped to its inverse by f , and denote by B the ball whose boundary passes through the endpoints of γ and is orthogonal to γ at them. Because γ is a geodesic, $B \subset D$. Because the circular complement of γ is a geodesic, $\overline{\mathbb{R}^n} \setminus \overline{B} \subset D$. Thus the boundary of D lies in a sphere, a case that we excluded. Therefore the points on this geodesic line are in F , so F intersects every component of D . Hence N is empty, so every point in D is fixed. \square

Acknowledgements. We would like to thank Fred Gehring for his efforts in promoting our collaboration, and for his encouragement and support.

REFERENCES

- [1] D. Barbilian, Einordnung von Lobatschewsky's Maßbestimmung in gewisse allgemeine Metrik der Jordanschen Bereiche, *Casopsis Matematiky a Fysiky* **64** (1934-35), 182–183.
- [2] A. Beardon, *The geometry of discrete groups*, Graduate Texts in Mathematics 91, Springer-Verlag, New York, 1995.

- [3] A. Beardon, The Apollonian metric of an open set in $\overline{\mathbb{R}^n}$, *Quasiconformal mappings and analysis* (eds. P. Duren, J. Heinonen, B. Osgood and B. Palka), Springer-Verlag, New York, 1998, pp. 91–108.
- [4] M. Berger, *Geometry I*, Springer-Verlag, Berlin, 1994.
- [5] M. Borovikova and Z. Ibragimov, Convex bodies of constant width and the Apollonian metric, in preparation.
- [6] W.-G. Boskoff, *Hyperbolic geometry and Barbilian spaces*, Series of Monographs in Advanced Mathematics, Hardronic Press Inc, Palm Harbor, FL, 1996.
- [7] W.-G. Boskoff, *Varietăți cu structură metrică Barbilian* (Romanian) [Manifolds with Barbilian metric structure] (Colecția Biblioteca de Matematică [Mathematics Library Collection], Ex Ponto, Editura, Constanța, 2002).
- [8] G. D. Chakerian and H. Groemer, Convex bodies of constant width, *Convexity and its applications* (eds. P. Gruber and J. Wills), Birkhäuser, Basel, 1983, pp. 49–96.
- [9] F. W. Gehring and K. Hag, The Apollonian metric and quasiconformal mappings, *In the tradition of Ahlfors and Bers* (eds. I. Kra and B. Maskit), Contemp. Math. 256, AMS, Providence, 2000, pp. 143–163.
- [10] F. W. Gehring and B. P. Palka, Quasiconformally homogeneous domain, *J. Anal. Math.* **30** (1976), 172–199.
- [11] P. A. Hästö, The Apollonian metric: uniformity and quasiconvexity, *Ann. Acad. Sci. Fenn. Math.* **28** (2003), 385–414.
- [12] P. A. Hästö, The Apollonian inner metric, *Comm. Anal. Geom.* **12** (2004), no. 4, 927–947.
- [13] P. A. Hästö, The Apollonian metric: the comparison property, bilipchitz mappings and thick sets, preprint (2005). Available at www.helsinki.fi/hasto/pp/.
- [14] P. Hästö and H. Lindén, Isometries of the half-apollonian metric, *Complex Var. Theory Appl.* **49** (2004), 405–415.
- [15] Z. Ibragimov, The Apollonian metric, Sets of constant width and Möbius modulus of ring domains, Ph.D. Thesis, University of Michigan, 2002.
- [16] Z. Ibragimov, On the Apollonian metric of domains in $\overline{\mathbb{R}^n}$, *Complex Var. Theory Appl.* **48** (2003), 837–855.
- [17] Z. Ibragimov, Conformality of the Apollonian metric, *Comput. Methods Funct. Theory* **3** (2003), 397–411.
- [18] P. J. Kelly, Barbilian geometry and the Poincaré model, *Amer. Math. Monthly* **61** (1954), 311–319.
- [19] A. G. Rhodes, An upper bound for the hyperbolic metric of a convex domain, *Bull. London Math. Soc.* **29** (1997), 592–594.
- [20] P. Seittenranta, Möbius-invariant metrics, *Math. Proc. Cambridge Philos. Soc.* **125** (1999), 511–533.