Check for
updates

Intrinsic Directions, Orthogonality, and Distinguished Geodesics in the Symmetrized Bidisc

Jim Agler[1] · Zinaida Lykova[2] · N. J. Young[2,3]

Abstract

The symmetrized bidisc

$$G \stackrel{\text{def}}{=} \{(z + w, zw) : |z| < 1, \quad |w| < 1\},$$

under the Carathéodory metric, is a complex Finsler space of cohomogeneity 1 in which the geodesics, both real and complex, enjoy a rich geometry. As a Finsler manifold, G does not admit a natural notion of angle, but we nevertheless show that there *is* a notion of orthogonality. The complex tangent bundle TG splits naturally into the direct sum of two line bundles, which we call the *sharp* and *flat* bundles, and which are geometrically defined and therefore covariant under automorphisms of G. Through every point of G, there is a unique complex geodesic of G in the flat direction, having the form

$$F^\beta \stackrel{\text{def}}{=} \{(\beta + \bar\beta z, z) \ : z \in \mathbb{D}\}$$

for some $\beta \in \mathbb{D}$, and called a *flat geodesic*. We say that a complex geodesic D *is orthogonal* to a flat geodesic F if D meets F at a point λ and the complex tangent space $T_\lambda D$ at λ is in the sharp direction at λ. We prove that a geodesic D has the closest point property with respect to a flat geodesic F if and only if D is orthogonal to F in the above sense. Moreover, G is foliated by the geodesics in G that are orthogonal to a fixed flat geodesic F.

Keywords Complex Finsler space · Symmetrized bidisc · Lempert domain · Geodesic · Orthogonality

Mathematics Subject Classification Primary: 32A07 · 53C22 · 54C15 · 47A57 · 32F45; Secondary: 47A25 · 30E05

Extended author information available on the last page of the article

Contents

1 Introduction

This paper concerns *geodesics* in the symmetrized bidisc G, the domain in $\mathbb{C}^2$ defined in the abstract. There are two different notions of geodesic, to wit, *complex* and *real*. A *complex geodesic* in G is defined to be the range of a holomorphic map $h : \mathbb{D} \to G$ which is isometric with respect to the Poincaré metric on the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and the hyperbolic metric on G. It is a fact that such an h is an isometry if and only if h has a holomorphic left inverse (see, for example, [11]). *Real geodesics* are paths which locally minimize lengths, in the usual sense of real differential geometry or metric geometry (see [10, Definition 2.5.27]). When we speak of a 'geodesic' in G, without qualification, we shall mean a *complex* geodesic.

We study the geometry of G as a hyperbolic complex space in the sense of Kobayashi [13]. It is well known that on G the *Carathéodory pseudometric* $|\delta|_{\mathrm{car}}$ and *Kobayashi pseudometric* $|\delta|_{\mathrm{kob}}$, where δ is a datum[1] in G, are in fact metrics, and moreover, they coincide [1,11]. Thus, there is a natural metric d_G on G, given, for a datum $\delta = (\lambda, \mu)$ in G, by

$$d_G(\delta) = |\delta|_{\mathrm{kob}} = |\delta|_{\mathrm{car}}.$$

We call d_G the *hyperbolic metric* on G.

---

[1] That is, either a pair of points in G or an element of the complex tangent bundle of G

A fundamental fact of the complex geometry of G is that, for every point $\lambda \in G$ and every one-dimensional subspace $\mathcal{M} \subseteq \mathbb{C}^2$, there exists a *unique* complex geodesic D in G such that $\lambda \in D$ and $T_\lambda D = \mathcal{M}$, where $T_\lambda D$ denotes the complex tangent space to D at the point λ. Thus, there is a natural one-to-one correspondence between the geodesics in G passing through λ and the points of CP^1, the *complex projective line*. We regard CP^1 as the set of one-dimensional subspaces of $\mathbb{C}^2$ and refer to its elements as *directions*. We shall exploit this correspondence to relate the local and global hyperbolic geometries of G.

As there are qualitative differences among the complex geodesics of G, so are there differences among the directions at a point. To illustrate this anisotropy, we first note that there exists a unique complex geodesic $\mathcal{R}$ in G that is invariant under all automorphisms of G.

Referred to as the *royal* variety, $\mathcal{R}$ is defined to be the set

$$\mathcal{R} = \{(2z, z^2), z \in \mathbb{D}\} = \{(s, p) \in G : s^2 = 4p\}.$$

Moreover, for every $\lambda \in \mathcal{R}$, there exists a unique geodesic F_λ in G having a nontrivial stabilizer in $\mathrm{Aut}\, G$, the group of biholomorphic self-maps of G, and such that

$$F_\lambda \cap \mathcal{R} = \{\lambda\}.$$

Thus, the set $\{F_\lambda : \lambda \in \mathcal{R}\}$ is a hyperbolically identifiable class of geodesics. We call them *flat* geodesics. They have the form

$$F^\beta \stackrel{\mathrm{def}}{=} \{(\beta + \bar{\beta}z, z) : z \in \mathbb{D}\} \tag{1.1}$$

for some $\beta \in \mathbb{D}$. Furthermore, through each point $\lambda \in G$, there passes a unique flat geodesic. Consequently, at each point $\lambda \in G$, there exists a unique direction $\flat_\lambda \in \mathrm{CP}^1$ with the property that

$$T_\lambda F = \flat_\lambda,$$

where F is the flat geodesic passing through λ. We refer to $\flat_\lambda$ as the *flat direction at* λ. The flat direction in G is covariant under automorphisms of G, in the sense that for every $\lambda \in G$ and $\gamma \in \mathrm{Aut}\, G$,

$$\flat_{\gamma(\lambda)} = \gamma'(\lambda)\flat_\lambda.$$

Another covariant direction is the *sharp* direction [5, Sect. 1.4]. For a point $\lambda \in G$, let $\mathrm{Orb}(\lambda)$ denote the orbit of λ under the action of $\mathrm{Aut}\, G$. In [5, Theorem 1.6], the authors showed that if $\lambda \in \mathcal{R}$, then $\mathrm{Orb}(\lambda) = \mathcal{R}$, and if $\lambda \in G \setminus \mathcal{R}$, then $\mathrm{Orb}(\lambda)$ is a smooth properly embedded 3-dimensional real manifold in G. Noting that, in both cases, $T_\lambda \mathrm{Orb}(\lambda)$ contains a unique one-dimensional complex subspace $\mathcal{M}$, we may define an intrinsic direction $\sharp_\lambda$, the *sharp direction at* λ, by $\sharp_\lambda \stackrel{\mathrm{def}}{=} \mathcal{M}$. Observe that,

for $\lambda \in \mathcal{R}$, since $\text{Orb}(\lambda) = \mathcal{R}$,

$$\sharp_\lambda = T_\lambda \, \text{Orb}(\lambda) = T_\lambda \mathcal{R}. \tag{1.2}$$

In this paper, besides the sharp and flat directions, we shall encounter other special directions at a point. We now describe the results of the paper.

If F is a flat geodesic in G and D is a general geodesic in G, we say that D *is orthogonal to* F if D meets F at a point λ and D points in the sharp direction at λ, that is, $T_\lambda D = \sharp_\lambda$. The motivation for this terminology is the following result (Corollary 7.11).

Theorem 1.1 *If F is a flat geodesic in G and D is an arbitrary geodesic in G, then D is orthogonal to F if and only if D meets F at a point λ_0 and, for some point $\mu \in D \setminus \{\lambda_0\}$ (equivalently, for every point $\mu \in D \setminus \{\lambda_0\}$),*

$$d_G(\lambda_0, \mu) = \inf_{\lambda \in F} d_G(\lambda, \mu).$$

The royal geodesic meets any flat geodesic in a unique point. Therefore, by the above definition and Eq. (1.2), $\mathcal{R}$ is orthogonal to every flat geodesic. If $D \neq \mathcal{R}$, then for D to be orthogonal to a flat geodesic, it must be the case that D is of a very special type which we now describe (see Corollary 5.5 and [6, Lemma 9.8]). We say that a geodesic D is *purely balanced* if $D^- \cap \mathcal{R}^-$ consists of exactly two points ξ_1 and ξ_2 both of which lie in $\partial \mathcal{R} = \{(2z, z^2) : z \in \mathbb{T}\}^2$. Here $\mathbb{T}$ denote the unit circle $\{z \in \mathbb{C} : |z| = 1\}$. The points ξ_1 and ξ_2 are referred to as the *royal points of* D. In Sect. 4 of this paper, we describe the purely balanced geodesics using properties of hyperbolic automorphisms of the disc.[3] This description is used to show that the purely balanced directions at a point in $G \setminus \mathcal{R}$ form a simple smooth curve in CP^1 connecting two 'exceptional' directions and containing the sharp direction as its midpoint (cf. Corollary 4.18).

We further exploit the description of the purely balanced geodesics to prove the following result in Sect. 5 of the paper (Theorem 5.10).

Theorem 1.2 *If F is a flat geodesic in G, then G is foliated by the geodesics in G that are orthogonal to F.*

As a corollary of Theorems 1.1 and 1.2 we obtain in Corollary 7.10 the following result.

Corollary 1.3 *If F is a flat geodesic in G and $\mu \in G$, then there exists a unique point $\lambda_0 \in F$ such that*

$$d_G(\lambda_0, \mu) = \inf_{\lambda \in F} d_G(\lambda, \mu).$$

---

[2] The closure and boundary symbols relate to D and $\mathcal{R}$ as subsets of $\mathbb{C}^2$.

[3] $m \in \text{Aut } \mathbb{D}$ is said to be *hyperbolic* if m has two distinct fixed points on $\mathbb{T}$.

[6, Lemma 9.8] contains the statement that the shortest distance from μ to F is always attained; we now see that it is attained at a unique point of F.

In Sect. 6, we turn the tables—we fix a purely balanced geodesic D and study the flat geodesics F such that D is orthogonal to F. If D is a geodesic, we let $\sharp(D)$ denote the set of λ in D such that D is orthogonal to the flat geodesic passing through λ (that is, $T_\lambda D = \sharp_\lambda$.) We say that a curve Ξ is a *simple real geodesic in G* if $\Xi = h(C)$, where $h : \mathbb{D} \to G$ is an isometry when $\mathbb{D}$ and G are equipped with their respective hyperbolic metrics, and C is a real hyperbolic geodesic[4] in $\mathbb{D}$. Each purely balanced geodesic D contains a unique real geodesic Ξ_D such that the endpoints of Ξ_D are the royal points of D (cf. Proposition 6.3).

Theorem 1.4 *If D is a purely balanced geodesic in G, then $\sharp(D) = \Xi_D$, that is, the set of $\lambda \in D$ such that $T_\lambda D = \sharp_\lambda$ is the simple real geodesic in G whose endpoints are the two royal points of D.*

As a corollary of Theorem 1.4, we obtain in Sect. 6 the following result.

Theorem 1.5 (cf. Theorem 6.8) *$G \setminus \mathcal{R}$ is foliated by the simple real geodesics in G whose endpoints lie in $\partial \mathcal{R}$.*

Another special direction at a point P of G, this time in the *real* tangent bundle of G, we call the *distinguished* direction. It is tangent to a special real geodesic γ through P with the property that, for every pair of points on γ, there are at least two inequivalent solutions of the corresponding Carathéodory extremal problem. These distinguished geodesics foliate $G \setminus \mathcal{R}$, where $\mathcal{R}$ is the 'royal variety' $\{(2z, z^2) : |z| < 1\}$.

The theory of the symmetrized bidisc, and cognate domains like the tetra-block, has been extensively developed over the last 20 years by numerous authors. We shall require some results from this theory, many of which can be found in [12] and [6, Appendix A]. These domains have a rich complex geometry and function theory, as well as applications to operator theory: see, besides many other papers, [4,7–9,12,14–16].

2 The Hyperbolic Metric on a Lempert Domain

In this section, we describe our terminology for the Carathéodory and Kobayashi extremal problems on an open set U in $\mathbb{C}^n$ and introduce the hyperbolic metric on a Lempert domain.

For open sets $U \subseteq \mathbb{C}^{n_1}$ and $V \subseteq \mathbb{C}^{n_2}$ we denote by $V(U)$ the set of holomorphic mappings from U into V.

If U is an open set in $\mathbb{C}^n$, then by a *datum in U* we mean an ordered pair δ where either δ is *discrete*, that is, has the form

$$\delta = (s_1, s_2)$$

where $s_1, s_2 \in U$, or δ is *infinitesimal*, that is, has the form

$$\delta = (s, v)$$

---

[4] That is, C is either a line segment or a circular arc in $\mathbb{D}$ that intersects $\mathbb{T}$ orthogonally.

where $s \in U$ and $v \in \mathbb{C}^n$.

If δ is a datum, we say that δ is *degenerate* if either δ is discrete and $s_1 = s_2$ or δ is infinitesimal and $v = 0$. Otherwise, we say that δ is *nondegenerate*.

An infinitesimal datum in U is the same thing as a point of the complex tangent bundle TU of U.

For $F \in \Omega(U)$, $s \in U$, and $v \in \mathbb{C}^{n_1}$, the directional derivative $D_v F(s) \in \mathbb{C}^{n_2}$ is defined by

$$D_v F(s) = \lim_{z \to 0} \frac{F(s + zv) - F(s)}{z}.$$

If U and Ω are domains, $F \in \Omega(U)$, and δ is a datum in U, we define a datum $F(\delta)$ in Ω by

$$F(\delta) = (F(s_1), F(s_2))$$

when $\delta = (s_1, s_2)$ is discrete and by

$$F(\delta) = (F(s), D_v F(s))$$

when $\delta = (s, v)$ is infinitesimal.

For any datum δ in $\mathbb{D}$, we define $|\delta|$ to be the Poincaré distance or metric at δ in the discrete or infinitesimal case respectably, that is

$$|\delta| = \tanh^{-1} \left| \frac{z_1 - z_2}{1 - \bar{z}_2 z_1} \right|$$

when $\delta = (z_1, z_2)$ is discrete[5], and by

$$|\delta| = \frac{|v|}{1 - |z|^2}$$

when $\delta = (z, v)$ is infinitesimal. See [17] for terminology and theory in several complex variables.

The Carathéodory extremal problem. *For a domain U in $\mathbb{C}^n$ and a nondegenerate datum δ in U, compute the quantity $|\delta|_{car}$ defined by*

$$|\delta|_{car} = \sup_{F \in \mathbb{D}(U)} |F(\delta)|. \tag{2.1}$$

We shall refer to this problem as Car δ and will say that C *solves* Car δ if $C \in \mathbb{D}(U)$ and

$$|\delta|_{car} = |C(\delta)|.$$

---

[5] In [6, Chap. 3], we used a different notation in that we omitted $\tanh^{-1}$; this makes no essential difference in the present context, but ensures that the Carathéodory pseudodistance is the inner pseudodistance determined by the Carathéodory pseudometric [12], and similarly for the Kobayashi pseudodistance.

$| \cdot |_{\text{car}}$ is a pseudometric on U, called the *Carathéodory pseudometric* in the case of infinitesimal δ and the *Carathéodory pseudodistance* in the case of discrete datums δ. It is easy to see with the aid of Montel's theorem that, for every nondegenerate datum δ in U, there does exist $C \in \mathbb{D}(U)$ which solves Car δ. Such a C is called a *Carathéodory extremal function* for δ.

Definition 2.1 We say that a domain U in $\mathbb{C}^n$ is *weakly hyperbolic* if $|\delta|_{\text{car}} > 0$ for every nondegenerate datum δ in U. Equivalently, for every nondegenerate datum δ in U, there exists a bounded holomorphic function F on U such that $F(\delta)$ is a nondegenerate datum in $\mathbb{C}$.

The Kobayashi extremal problem. *For a domain U in $\mathbb{C}^n$ and a nondegenerate datum δ in U, compute the quantity $|\delta|_{\text{kob}}$ defined by*

$$|\delta|_{\text{kob}} = \inf_{\substack{f \in U(\mathbb{D}) \\ f(\zeta) = \delta}} |\zeta|. \tag{2.2}$$

We shall refer to this problem as Kob δ and will say that k *solves* Kob δ if $k \in U(\mathbb{D})$ and there exists a datum ζ in $\mathbb{D}$ such that $k(\zeta) = \delta$ and

$$|\delta|_{\text{kob}} = |\zeta|.$$

On infinitesimal datums $| \cdot |_{\text{kob}}$ is a pseudometric, called the *Kobayashi pseudometric* or the *Kobayashi–Royden pseudometric* [12, Chap. 3]. The quantity $| \cdot |_{\text{kob}}$ is not necessarily a pseudodistance on discrete datums (it can fail to satisfy the triangle inequality). The *Kobayashi pseudodistance* on U is defined to be the largest pseudodistance on U majorized by $| \cdot |_{\text{kob}}$.

Note that the infimum in the definition (2.2) of $|\delta|_{\text{kob}}$ is attained if U is a taut domain, where U is said to be *taut* if $U(\mathbb{D})$ is a normal family. In particular, $|\delta|_{\text{kob}}$ is attained when $U = G$ [12, Sect. 3.2]. Any function which solves Kob δ is called a *Kobayashi extremal function for δ*.

The Kobayashi and Carathéodory pseudometrics are *invariant*, that is, they are invariant under automorphisms of G; see [12] for an up-to-date account of such pseudometrics.

Let U be a domain in $\mathbb{C}^n$ and δ a nondegenerate datum in U. The solutions to Car δ and Kob δ are never unique, for if m is a Möbius transformation of $\mathbb{D}$, then $m \circ C$ solves Car δ whenever C solves Car δ and $f \circ m$ solves Kob δ whenever f solves Kob δ. This suggests the following definition.

Definition 2.2 Let U be a domain in $\mathbb{C}^n$ and let δ be a nondegenerate datum in U. We say that *the solution to Car δ is essentially unique*, if whenever F_1 and F_2 solve Car δ there exists a Möbius transformation m of $\mathbb{D}$ such that $F_2 = m \circ F_1$. We say that *the solution to Kob δ is essentially unique* if the infimum in Eq. (2.2) is attained and, whenever f_1 and f_2 solve Car δ there exists a Möbius transformation m of $\mathbb{D}$ such that $f_2 = f_1 \circ m$.

In honor of Lempert's seminal theorem [18], we adopt the following definition.

Definition 2.3 A domain U in $\mathbb{C}^n$ is a *Lempert domain* if

(1) U is weakly hyperbolic,
(2) U is taut, and
(3) $|\delta|_{car} = |\delta|_{kob}$ for every nondegenerate datum δ in U.

Thus, for a Lempert domain U in $\mathbb{C}^n$, there is a natural metric d_U on U, given, for a nondegenerate infinitesimal datum δ in U, by

$$d_U(\delta) = |\delta|_{kob} = |\delta|_{car}.$$

The metric d_U is called the *hyperbolic metric* on U.
The Hyperbolic Metric on the Symmetrized Bidisc The *symmetrization map*

$$\pi : \mathbb{C}^2 \to \mathbb{C}^2$$

is defined by

$$\pi\big((z_1, z_2)\big) = (z_1 + z_2, z_1 z_2) \qquad \text{for all } z_1, z_2 \in \mathbb{C}.$$

Thus, the symmetrized bidisc G is $\pi(\mathbb{D}^2)$.
   We adopt the coordinates $s = z_1 + z_2$ and $p = z_1 z_2$. For $s, p \in \mathbb{C}$,

$$(s, p) \in G \text{ if and only if } |s - \bar{s}p| < 1 - |p|^2 \tag{2.3}$$

(for example [12, Lemma 7.1.3]).
   The symmetrized bidisc G is a Lempert domain in $\mathbb{C}^2$ with the hyperbolic metric d_G [1].

3 Automorphisms, Complex Geodesics and Directions in G

3.1 The Automorphism Group of G

For a domain Ω in $\mathbb{C}^d$, an *automorphism* of Ω is defined to be a biholomorphic self-map of Ω. The group of all automorphisms of Ω under composition will be denoted by Aut Ω. Here is a description of Aut G in terms of Aut $\mathbb{D}$.

Proposition 3.1 *For every $b \in$ Aut $\mathbb{D}$, there is a unique automorphism γ_b of G satisfying*

$$\gamma_b(\pi(z_1, z_2)) = \pi(b(z_1), b(z_2)) \qquad \text{for all } z_1, z_2 \in \mathbb{D}. \tag{3.1}$$

Furthermore, the map $f :$ Aut $\mathbb{D} \to$ Aut G given by

$$f(b) = \gamma_b \qquad \text{for all } b \in \text{Aut } \mathbb{D}$$

is a continuous isomorphism with respect to the compact-open topologies on $\mathrm{Aut}\,\mathbb{D}$
and $\mathrm{Aut}\,G$.

A proof can be found in [12, Sect. 7.1] or [3, Theorem 4.1].

The following corollary of Proposition 3.1 is an immediate consequence of the fact that every automorphism of $\mathbb{D}$ extends to be analytic in a neighborhood of $\mathbb{D}^-$.

Corollary 3.2 (1) *Every automorphism γ of G extends in a unique way to a self-homeomorphism $\tilde{\gamma}$ of the closure of G in $\mathbb{C}^2$;*
(2) $\tilde{\gamma}(\mathcal{R}^-) = \mathcal{R}^-$;
(3) *the classes of flat and purely balanced geodesics in G are invariant under the automorphisms of G.*

3.2 The Complex Geodesics in *G*

By a *complex geodesic* in G we mean a subset $D \subset G$ such that there exists a hyperbolic isometry $h : \mathbb{D} \to G$ such that $D = h(\mathbb{D})$. It is known that through any two points of G there passes a unique geodesic (see [2, Theorem 0.3]). We note that $h : \mathbb{D} \to G$ is a hyperbolic isometry if and only if h is holomorphic and there exists a holomorphic map $\Phi : G \to \mathbb{D}$ such that[6] $\Phi \circ h = \mathrm{id}_{\mathbb{D}}$. Moreover, when $D = h(\mathbb{D})$ is a geodesic, h is rational and extends to be holomorphic on a neighborhood of $\mathbb{D}^-$, so that, in particular, $D^- = h(\mathbb{D}^-)$.

There are qualitative differences between the geodesics of G. Indeed, in [6, Chap. 7] five distinct types of geodesic are identified, namely *royal, flat, purely unbalanced, purely balanced,* and *exceptional*. Each of the types can be characterized in terms of qualitative properties of the Carathéodory extremal problem associated with the geodesic and in a variety of purely geometric ways as well.

Here, we adopt a description of the five types of geodesics in terms of the geometry of the distinguished boundary[7] of G, which can be shown to be the set in $\mathbb{C}^2$ defined by

$$M = \pi(\mathbb{T}^2).$$

Topologically, M is a Möbius band. The *edge* E of the Möbius band is the set in $\mathbb{C}^2$

$$E = \pi(\{(z, z) \ : \ z \in \mathbb{T}\}).$$

Note that $E = \partial \mathcal{R}$.

---

[6] Here $\mathrm{id}_{\mathbb{D}}$ denotes the automorphism defined by $\mathrm{id}_{\mathbb{D}}(z) = z, \ z \in \mathbb{D}$.

[7] That is, the smallest closed subset M of G^- with the property that every holomorphic function defined on a neighborhood of G^- attains its maximum over G^- at a point in M.

3.3 The Five Types of Geodesic

We summarize the definitions of the types of geodesic introduced in [6] and state a few of their properties. In that paper [6], the types were defined in Sect. 3.3, but here we shall rather use a geometric characterization of them given in [6, Chap. 7].

(1) The *royal* geodesic is the set

$$\mathcal{R} = \pi(\{(z, z) : z \in \mathbb{D}\}).$$

It is the only geodesic left invariant by every automorphism in Aut G. Furthermore[8],

$$\mathcal{R}^- \cap \mathrm{M} = \mathrm{E}.$$

Note also that $\mathcal{R}^-$ is the disjoint union of $\mathcal{R}$ and E.

If D is a geodesic we define the *royal points* of D to be the elements of the set $D^- \cap \mathcal{R}^-$. Since $\mathcal{R}^- = \mathcal{R} \cup \mathrm{E}$, each royal point must either be in $\mathcal{R}$ or in E.

(2) A geodesic D is said to be *flat* if $D^- \cap \mathcal{R}^-$ consists of a single point and that point lies[9] in $\mathcal{R}$. Thus, a geodesic D is flat if it has a unique royal point and that point lies in $\mathcal{R}$ (rather than E). The nomenclature "flat geodesic" reflects the fact that a geodesic D is flat if and only if D lies in a set of the form $\lambda + \mathcal{M}$ where $\lambda \in G$ and $\mathcal{M}$ is a one-dimensional complex subspace of $\mathbb{C}^2$. There exists a unique flat geodesic passing through each point of G, or in other words, the flat geodesics foliate G. The royal geodesic is the only geodesic in G that meets every flat geodesic.

(3) A geodesic D is said to be *purely balanced* if $D^- \cap \mathcal{R}^-$ consists of exactly two points both of which lie in E, that is, D has exactly two royal points, and they lie in E. We shall give a concrete formula for the general purely balanced geodesic using hyperbolic automorphisms of $\mathbb{D}$ in Sect. 4.

(4) A geodesic D is said to be *purely unbalanced* if $D^- \cap \mathcal{R}^-$ consists of exactly two points, one of which lies in $\mathcal{R}$ and one of which lies in E, that is, D has exactly two royal points, one of which lies in $\mathcal{R}$ and one of which lies in E. Generically, the geodesic that passes through two distinct points in G is purely unbalanced.

(5) A geodesic D is said to be *exceptional* if $D^- \cap \mathcal{R}^-$ consists of exactly one point and that point lies in E, that is, D has exactly one royal point, and that point lies in E.

3.4 Directions

We shall denote a direction at a point in G (that is, an element of CP^1) by

$$v\mathbb{C} = \{zv : z \in \mathbb{C}\} \qquad \text{where } v \in \mathbb{C}^2 \setminus \{0\}.$$

---

[8] Indeed, $\mathcal{R}^-$ is the polynomially convex hull of E.

[9] Equivalently, a geodesic D is flat if and only if $D^- \cap \mathrm{E} = \emptyset$.

The following pleasing state of affairs is central to this paper (see [6, Theorem 4.6]).

Theorem 3.3 *For every point λ in G and every nonzero vector v in $\mathbb{C}^2$, there exists a unique geodesic D^v in G such that*

$$\lambda \in D^v \ \text{ and } \ v \in T_\lambda D^v.$$

Furthermore, if v and w are any two nonzero vectors in $\mathbb{C}^2$, $D^v = D^w$ if and only if $v\mathbb{C} = w\mathbb{C}$. In particular,

$$v\mathbb{C} \mapsto D^v \tag{3.2}$$

is a well-defined bijection between directions in CP^1 and geodesics in G that pass through λ.

Evidently, Theorem 3.3 implies that just as there are five qualitatively distinct types of geodesic, so are there five qualitatively distinct types of direction at a point. Let us say that a direction $v\mathbb{C}$ is of a particular type if the geodesic D^v that corresponds to $v\mathbb{C}$ via the map (3.2) is of that same type.

3.4.1 The Flat Direction

As noted above, through each point of G there passes a unique flat geodesic in G. Hence, for each $\lambda \in G$ there is a unique direction $\flat_\lambda \in \mathrm{CP}^1$ defined by the following procedure.

(1) Let D be the unique flat geodesic in G passing through λ.
(2) Choose any nonzero vector v in $T_\lambda D$.
(3) Let $\flat_\lambda = v\mathbb{C}$.

We refer to $\flat_\lambda$ as the *flat direction at λ* and say that a nonzero vector $v \in \mathbb{C}^2$ *points in the flat direction at λ* if $v \in \flat_\lambda$.

3.4.2 The Sharp Direction

In [5, Theorem 1.6], the authors showed that if $\lambda \in G$ and $\mathrm{Orb}(\lambda)$ denotes the orbit of λ under $\mathrm{Aut}\, G$, then

(1) $\mathrm{Orb}(\lambda) = \mathcal{R}$ if $\lambda \in \mathcal{R}$, and
(2) $\mathrm{Orb}(\lambda)$ is a smooth properly embedded 3-dimensional real manifold in G if $\lambda \notin \mathcal{R}$.

As a consequence of this result, for each $\lambda \in G$, we may define $\sharp_\lambda \in \mathrm{CP}^1$, by the following procedure.

(1) If $\lambda \in \mathcal{R}$, then $\sharp_\lambda = v\mathbb{C}$, where v is any nonzero vector in $T_\lambda \mathcal{R}$.
(2) Otherwise, if $\lambda \notin \mathcal{R}$, let $\sharp_\lambda = v\mathbb{C}$, where v is any vector such that both v and iv are in the real tangent space to $\mathrm{Orb}(\lambda)$ at λ.

We refer to $\sharp_\lambda$ as the *sharp direction at λ* and we say that a nonzero vector $v \in \mathbb{C}^2$ *points in the sharp direction at λ* if $v \in \sharp_\lambda$. Evidently, with this language, if $\lambda \in G$ and v is a nonzero vector in $\mathbb{C}^2$, then v points in the sharp direction at λ if and only if v is in the unique nonzero complex subspace of the real tangent space to $\mathrm{Orb}(\lambda)$ at λ.

3.4.3 Other Directions

For each point λ of G, the previous two subsections constructed a unique pair of directions $\flat_\lambda$ and $\sharp_\lambda$. These directions, being defined geometrically, are covariant in the sense that if γ is an automorphism of G, then

$$\flat_{\gamma(\lambda)} = \gamma'(\lambda)\,\flat_\lambda \quad \text{and} \quad \sharp_{\gamma(\lambda)} = \gamma'(\lambda)\,\sharp_\lambda \quad \text{for all } \lambda \in G.$$

The consideration of covariant directions other than the flat and sharp directions is complicated by issues of both existence and uniqueness. At a point $\lambda \in G$, there may not be any directions of a given type. Alternatively, there may be multiple directions of a given type. We briefly savor the low-hanging fruit regarding these issues.

(1) As the royal geodesic $\mathcal{R}$ is unique, if $\lambda \in G$, then there exists a royal geodesic passing through λ if and only if $\lambda \in \mathcal{R}$. Therefore, there exists a 'royal direction' at λ only when $\lambda \in \mathcal{R}$. However, in that case, $T\mathcal{R}_\lambda = \sharp_\lambda$, that is, a 'royal direction' exists only if it is the sharp direction. For this reason, we shall not henceforth use the term royal direction.

(2) According to our definition, a direction $v\mathbb{C}$ is purely balanced at $\lambda \in G$ if the unique geodesic D passing through λ satisfying $T_\lambda D = v\mathbb{C}$ is purely balanced. Notice from item (4) above in the discussion of the five types of geodesic that no purely balanced geodesic can meet $\mathcal{R}$. Therefore, if $\lambda \in \mathcal{R}$, then there are no purely balanced directions at λ.

(3) On the other hand, if $\lambda \in G \setminus \mathcal{R}$, it turns out that there is a one-parameter family of purely balanced directions at λ, see (7).

(4) A direction $v\mathbb{C}$ is *exceptional at* $\lambda \in G$ if the unique geodesic D passing through λ satisfying $T_\lambda D = v\mathbb{C}$ is exceptional. Notice from item (5) above in the discussion of the five types of geodesics that no exceptional geodesic can meet $\mathcal{R}$. Therefore, if $\lambda \in \mathcal{R}$, then there are no exceptional directions at λ.

(5) On the other hand, if $\lambda \in G \setminus \mathcal{R}$, it turns out that there are exactly two exceptional[10] directions $\natural^1_\lambda$ and $\natural^2_\lambda$ at λ. We let $\natural_\lambda$ denote the set of these two directions, that is, $\natural_\lambda = \{\natural^1_\lambda, \natural^2_\lambda\}$, see (7).

(6) Strictly speaking, the exceptional directions are not covariant. However, the *pair* of exceptional direction is covariant: if $\lambda \in G \setminus \mathcal{R}$ and $\gamma \in \text{Aut}\, G$, then

$$\natural_{\gamma(\lambda)} = \gamma'(\lambda)(\natural_\lambda).$$

(7) If $\lambda \in G \setminus \mathcal{R}$, then the purely balanced directions at λ form a simple smooth curve connecting $\natural^1_\lambda$ and $\natural^2_\lambda$ in CP^1 (cf. Corollary 4.18).

(8) If $\lambda \in G$, we say a direction $v\mathbb{C}$ is *purely unbalanced* at λ if the unique geodesic D passing through λ satisfying $T_\lambda D = v\mathbb{C}$ is purely unbalanced.

(9) If $\lambda \in \mathcal{R}$ and $v\mathbb{C}$ is a direction at λ, then either $v\mathbb{C} = \flat_\lambda$, $v\mathbb{C} = \sharp_\lambda$, or $v\mathbb{C}$ is purely unbalanced.

(10) If $\lambda \in G \setminus \mathcal{R}$ and $v\mathbb{C}$ is a direction at λ, then either $v\mathbb{C} = \flat_\lambda$, $v\mathbb{C} \in \natural_\lambda$, $v\mathbb{C}$ is purely balanced, or $v\mathbb{C}$ is purely unbalanced or exceptional. The sharp direction

---

[10] The musical symbol $\natural$ is read 'natural'.

is the midpoint of the purely balanced curve alluded to in item (7) above (cf. Corollary 5.7).

4 Hyperbolic Automorphisms in $\mathbb{D}$ and Purely Balanced Geodesics in G

4.1 Hyperbolic Automorphisms of $\mathbb{D}$

If $\alpha \in \mathbb{D}$ we define $b_\alpha \in \mathrm{Aut}\,\mathbb{D}$ by

$$b_\alpha(z) = \frac{z - \alpha}{1 - \bar{\alpha}z}, \quad z \in \mathbb{D},$$

and when $\tau \in \mathbb{T}$, we define $r_\tau \in \mathrm{Aut}\,\mathbb{D}$ by

$$r_\tau(z) = \tau z, \quad z \in \mathbb{D}.$$

Each $m \in \mathrm{Aut}\,\mathbb{D}$ can be uniquely represented in the form

$$m = r_\tau \circ b_\alpha \tag{4.1}$$

for some $\alpha \in \mathbb{D}$ and $\tau \in \mathbb{T}$.

If $m \in \mathrm{Aut}\,\mathbb{D} \setminus \{\mathrm{id}_\mathbb{D}\}$ then m can be viewed as a continuous self-map of $\mathbb{D}^-$ and it is well known that exactly one of the following cases occurs.

  (i) m is *elliptic*, that is, m has exactly one fixed point in $\mathbb{D}^-$, which lies in $\mathbb{D}$.
 (ii) m is *parabolic*, that is, m has exactly one fixed point in $\mathbb{D}^-$, which lies in $\mathbb{T}$.
(iii) m is *hyperbolic*, that is, m has exactly two fixed points in $\mathbb{D}^-$, which lie in $\mathbb{T}$.

We record the following lemma which gives a well-known criterion for the type of an automorphism in terms of the parameters α and τ in its representation as in Eq. (4.1).

Lemma 4.1 *Let $m \in \mathrm{Aut}\,\mathbb{D} \setminus \{\mathrm{id}_\mathbb{D}\}$ and assume that $m = r_\tau \circ b_\alpha$ where $\alpha \in \mathbb{D}$ and $\tau \in \mathbb{T}$.*

  (i) *m is elliptic* $\Leftrightarrow$ $|\tau - 1| > 2|\alpha|$.
 (ii) *m is parabolic* $\Leftrightarrow$ $|\tau - 1| = 2|\alpha|$.
(iii) *m is hyperbolic* $\Leftrightarrow$ $|\tau - 1| < 2|\alpha|$.

Note that this lemma implies that b_{z_0} is hyperbolic for every $z_0 \in \mathbb{D} \setminus \{0\}$.

Lemma 4.2 *If $\alpha, \beta \in \mathbb{D}$ and $\beta \neq -\alpha$, then $b_\beta \circ b_\alpha$ is hyperbolic.*

Proof Clearly $b_\beta \circ b_\alpha \neq \mathrm{id}_\mathbb{D}$. By direct calculation,

$$(b_\beta \circ b_\alpha)(z) = \frac{1 + \bar{\alpha}\beta}{1 + \bar{\beta}\alpha} \frac{z - \frac{\alpha + \beta}{1 + \bar{\alpha}\beta}}{1 - \frac{\bar{\alpha} + \bar{\beta}}{1 + \bar{\beta}\alpha}z}.$$

Hence, Lemma 4.1 implies that $b_\beta \circ b_\alpha$ is hyperbolic if and only if

$$\left|\frac{1+\bar{\alpha}\beta}{1+\bar{\beta}\alpha} - 1\right| < 2\left|\frac{\alpha+\beta}{1+\bar{\alpha}\beta}\right|$$

which simplifies to

$$|\bar{\alpha}\beta - \bar{\beta}\alpha| < 2|\alpha + \beta|.$$

But, as $|\alpha|, |\beta| < 1$ and $|\alpha + \beta| \neq 0$,

$$|\bar{\alpha}\beta - \bar{\beta}\alpha| = |\bar{\alpha}(\alpha + \beta) - \beta(\bar{\alpha} + \bar{\beta})| \leq |\alpha||\alpha + \beta| + |\beta||\alpha + \beta| < 2|\alpha + \beta|. \qquad \square$$

In the following definition, we introduce a class of automorphisms that plays a special role in the study of orthogonality.

Definition 4.3 If $m \in \operatorname{Aut} \mathbb{D}$, we say m is *irrotational* if in the representation of m in Eq. (4.1), $\tau = 1$.

Remark 4.4 (1) Lemma 4.1 implies that if m is irrotational, then either $m = \operatorname{id}_\mathbb{D}$ or m is hyperbolic. Lemma 4.2 now implies that if m_1 and m_2 are irrotational then either $m_2 = m_1^{-1}$ or $m_2 \circ m_1$ is hyperbolic.
(2) Observe, by direct calculation, that for $m \in \operatorname{Aut} \mathbb{D}$,

$$m \text{ is irrotational} \quad \Leftrightarrow \quad m'(0) > 0.$$

The following lemma characterizes when an automorphism is irrotational in terms of its fixed points.

Lemma 4.5 *If $m \in \operatorname{Aut} \mathbb{D}$, then m is irrotational if and only if there exists $\eta \in \mathbb{T}$ such that $m(\eta) = \eta$ and $m(-\eta) = -\eta$.*

Proof First assume that $m = b_\alpha$ is irrotational. When $\alpha = 0$, *every* point in $\mathbb{T}$ is a fixed point of m and when $\alpha \neq 0$, by direct computation the fixed points of m are the roots of the equation $\eta^2 = \alpha/\bar{\alpha}$.

Conversely, assume that $\eta \in \mathbb{T}, m(\eta) = \eta$, and $m(-\eta) = -\eta$. By the representation for m given in Eq. (4.1), we have the equations

$$\tau \frac{\eta - \alpha}{1 - \bar{\alpha}\eta} = \eta \quad \text{and} \quad \tau \frac{-\eta - \alpha}{1 + \bar{\alpha}\eta} = -\eta, \tag{4.2}$$

which imply upon elimination of τ that $\eta^2 = \alpha/\bar{\alpha}$. By the first of Eq. (4.2),

$$\tau(\eta - \alpha) = \eta(1 - \bar{\alpha}\eta) = \eta - \alpha,$$

which implies that $\tau = 1$. $\qquad \square$

The following lemma examines a special point $\alpha \in \mathbb{D}$ that can be associated with an irrotational automorphism. This point α will play an essential role in our study of orthogonality in G.

Lemma 4.6 *Let $m \in \mathrm{Aut}\,\mathbb{D}$. The following are equivalent.*

(i) *m is irrotational;*
(ii) *there exists $\alpha \in \mathbb{D}$ such that $m(\alpha) = -\alpha$ and $m'(\alpha) = 1$;*
(iii) *there exists $\alpha \in \mathbb{D}$ such that $m = b_\alpha \circ b_\alpha$;*
(iv) *$-m$ is hyperbolic reflection about a point α in $\mathbb{D}$.*

Proof (i) $\Rightarrow$ (ii). Let $m = b_\beta$ where $\beta \in \mathbb{D}$. Lemma 4.1 implies that $-b_\beta$ is elliptic. Choose $\alpha \in \mathbb{D}$ such that $-b_\beta(\alpha) = \alpha$. We have

$$b'_\beta(\alpha) = \frac{1 - |\beta|^2}{(1 - \bar{\beta}\alpha)^2} = \frac{(1 - \bar{\beta}\alpha) + \bar{\beta}(\alpha - \beta)}{(1 - \bar{\beta}\alpha)^2}$$

$$= \frac{1}{1 - \bar{\beta}\alpha} + \frac{\bar{\beta}}{1 - \bar{\beta}\alpha} b_\beta(\alpha) = \frac{1}{1 - \bar{\beta}\alpha} + \frac{\bar{\beta}}{1 - \bar{\beta}\alpha}(-\alpha) = 1.$$

Hence (ii) holds.

(ii) $\Rightarrow$ (iii). One can show by Schur reduction that $m = b_\alpha \circ r_\tau \circ b_\alpha$ for some $\tau \in \mathbb{T}$. But by the chain rule,

$$1 = m'(\alpha) = b'_\alpha(0) \cdot \tau \cdot b'_\alpha(\alpha) = (1 - |\alpha|^2) \cdot \tau \cdot \frac{1}{1 - |\alpha|^2} = \tau.$$

(iii) $\Rightarrow$ (iv). By direct calculation, $-b_\alpha = b_{-\alpha} \circ r_{-1}$. Therefore, as $b_{-\alpha} = b_\alpha^{-1}$

$$-b_\alpha \circ b_\alpha = b_\alpha^{-1} \circ r_{-1} \circ b_\alpha.$$

Hence, as $b_\alpha^{-1} \circ r_{-1} \circ b_\alpha$ is hyperbolic reflection about α, (iv) holds.

(iv) $\Rightarrow$ (i). If $-m$ is hyperbolic reflection about α, then

$$m = -b_\alpha^{-1} \circ r_{-1} \circ b_\alpha = b_\alpha \circ b_\alpha = b_\beta$$

where

$$\beta = \frac{2\alpha}{1 + |\alpha|^2}. \qquad \square$$

Note that each of the α's which appear in Conditions (ii), (iii), and (iv) of Lemma 4.6 is equal and is uniquely determined by m. Also note that the lemma implies that $b_\alpha \mapsto b_\alpha \circ b_\alpha$ is an injective map defined on the irrotational automorphisms onto the irrotational automorphisms. In particular, each irrotational automorphism has a unique irrotational square root with respect to composition.

4.2 Purely Balanced Geodesics and Royal Points in G

If $m \in \text{Aut}\,\mathbb{D}$, then the formula,

$$h_m(z) = (z + m(z), zm(z)), \qquad z \in \mathbb{D}, \tag{4.3}$$

defines a mapping $h_m : \mathbb{D} \to G$. It is easy to see that, when m is elliptic, h_m is not injective, and thus cannot isometrically parametrize a geodesic. In all other cases, h_m does parametrize a geodesic which we denote by B_m. The essentials of the correspondence between the properties of m and B_m are described in the next proposition.

Proposition 4.7 *Let* $m \in \text{Aut}\,\mathbb{D}$.

  (i) $m = \text{id}_{\mathbb{D}} \quad \Leftrightarrow \quad B_m = \mathcal{R}$.
 (ii) *m is hyperbolic* $\quad \Leftrightarrow \quad B_m$ *is purely balanced*.
(iii) *m is parabolic* $\quad \Leftrightarrow \quad B_m$ *is exceptional*.

Proof This proposition is proved in [6, Theorem 7.8]. □

Lemma 4.8 *If* m, q *are hyperbolic automorphisms of* $\mathbb{D}$ *and* $B_m = B_q$ *then either* $q = m$ *or* $q = m^{-1}$.

Proof Let the fixed points of m, q be $\eta = \{\eta_1, \eta_2\}$ and $\zeta = \{\zeta_1, \zeta_2\}$, respectively. Then $h_q^{-1} \circ h_m$ is an automorphism, υ say, of $\mathbb{D}$. Thus, $h_m = h_q \circ \upsilon$, that is, for all $z \in \mathbb{D}$,

$$(z + m(z), zm(z)) = h_q(\upsilon(z)) = (\upsilon(z) + q \circ \upsilon(z), \upsilon(z)q \circ \upsilon(z)).$$

Therefore, for all z, either

$$\upsilon(z) = z \quad \text{and} \quad q \circ \upsilon(z) = m(z)$$

or

$$q \circ \upsilon(z) = z \quad \text{and} \quad \upsilon(z) = m(z).$$

Thus, either $\upsilon = \text{id}_{\mathbb{D}}$ and $q = m$, or $\upsilon = m$ and $q = m^{-1}$. □

For hyperbolic $m \in \text{Aut}\,\mathbb{D}$, there are important relationships between the fixed points of m and properties of the corresponding purely balanced geodesic B_m. We describe three of these in Lemma 4.9, Corollary 4.12, and Proposition 6.3.

Recall that a geodesic D was defined to be purely balanced if $D^- \cap \mathcal{R}^-$ consists of exactly two points, both of which lie in E.

Lemma 4.9 *If m is hyperbolic and η_1 and η_2 are the distinct fixed points of m in $\mathbb{T}$, then*

$$B_m^- \cap \mathcal{R}^- = \left\{(2\eta_1, \eta_1^2), (2\eta_2, \eta_2^2)\right\}.$$

Proof If $m(\eta) = \eta$, then $h_m(\eta) = (2\eta, \eta^2)$. Therefore,

$$(2\eta_1, \eta_1^2), (2\eta_2, \eta_2^2) \in B_m^- \cap \mathcal{R}^-.$$

Conversely, if $\lambda \in B_m^- \cap \mathcal{R}^-$, then since $\lambda \in B_m^-$, there exists $\zeta \in \mathbb{T}$ such that

$$\lambda = (\zeta + m(\zeta), \zeta m(\zeta)),$$

and since $\lambda \in \mathcal{R}^-$, there exists $\eta \in \mathbb{T}$ such that

$$\lambda = (2\eta, \eta^2).$$

These equations imply that

$$\zeta + m(\zeta) = 2\eta \quad \text{and} \quad \zeta m(\zeta) = \eta^2.$$

Hence, for all $x \in \mathbb{C}$,

$$
\begin{aligned}
(x - \zeta)(x - m(\zeta)) &= x^2 - (\zeta + m(\zeta))x + \zeta m(\zeta) \\
&= x^2 - 2\eta x + \eta^2 = (x - \eta)^2,
\end{aligned}
$$

which implies that $\zeta = \eta$ and so $m(\eta) = \eta$.   □

A second relationship between B_m and the fixed points of m involves the qualitative nature of the solutions to the Carathéodory extremal problem. We say, for a domain Ω, that a set $\mathcal{C}$ of holomorphic maps from Ω to $\mathbb{D}$ is a *universal set for the Carathéodory extremal problem on Ω* if, for every $\lambda \in \Omega$ and every nonzero vector $v \in T_\lambda \Omega$, there exists $\Phi \in \mathcal{C}$ such that the supremum in Eq. (2.1), when $U = \Omega$, is attained at $F = \Phi$.

For $\omega \in \mathbb{T}$, define a holomorphic function Φ_ω on G by the formula,

$$\Phi_\omega(s, p) = \frac{2\omega p - s}{2 - \omega s}, \qquad \text{for all } (s, p) \in G.$$

The following properties of Φ_ω, $\omega \in \mathbb{T}$, are established in [1, Theorem 2.1 and Corollary 4.3].

Proposition 4.10 *For every $\omega \in \mathbb{T}$, Φ_ω is a holomorphic map from G to $\mathbb{D}$. Furthermore, the set $\{\Phi_\omega : \omega \in \mathbb{T}\}$ is universal for the Carathéodory extremal problem on G.*

The following result gives a criterion for a geodesic in G to be purely balanced in terms of a qualitative property of solutions of the Carathéodory extremal problem. It is contained in [6, Theorem 7.1(iii)].

Proposition 4.11 *Let D be a geodesic in G, let $\lambda \in D$ and let v be a nonzero vector in $T_\lambda D$. The geodesic D is purely balanced if and only if there are two distinct points $\omega_1, \omega_2 \in \mathbb{T}$ such that the supremum in Eq. (2.1), when $U = G$, is attained at $F = \Phi_\omega$*

precisely when either $\omega = \omega_1$ or $\omega = \omega_2$. Furthermore, if D is purely balanced and ω_1, ω_2 are as in the preceding sentence, then the royal points of D are $(2\bar{\omega}_1, \bar{\omega}_1^2)$ and $(2\bar{\omega}_2, \bar{\omega}_2^2)$.

The second promised relationship between a purely balanced geodesic and its royal points is contained in the following statement.

Corollary 4.12 *Let m be a hyperbolic automorphism of $\mathbb{D}$, let η_1 and η_2 be the distinct fixed points of m in $\mathbb{T}$, let $\lambda \in B_m$ and let $v \in T_\lambda B_m$. Then Φ_ω is a solution of the Carathéodory extremal problem on G at the point λ in direction v if and only if either $\omega = \bar{\eta}_1$ or $\omega = \bar{\eta}_2$.*

4.3 Purely Balanced Geodesics in Standard Position in *G*

In this subsection, we set out a canonical form for purely balanced geodesics. We accomplish this goal by cleanly parametrizing the purely balanced geodesics that pass through a point of the form $(0, p)$ where $p < 0$.

Lemma 4.13 *Let B be a purely balanced (or exceptional) geodesic and let $\sigma \in (0, 1)$.*

$$(0, -\sigma^2) \in B$$

if and only if there exists $m \in \operatorname{Aut} \mathbb{D}$ satisfying

 (i) *m is hyperbolic (or parabolic, respectively),*
(ii) *$m(\sigma) = -\sigma$, and*
(iii) *$B = B_m$.*

Proof If conditions (i), (ii) and (iii) hold, then

$$(0, -\sigma^2) = \left(\sigma + (-\sigma), \sigma(-\sigma)\right) = h_m(\sigma) \in B_m = B.$$

Furthermore, Proposition 4.7 implies that B is purely balanced (or exceptional, respectively).

Conversely, assume that B is purely balanced (or exceptional) and $(0, -\sigma^2) \in B$. Using Proposition 4.7, we may choose a hyperbolic (or exceptional, respectively) $b \in \operatorname{Aut} \mathbb{D}$ such that $B = B_b$. Choose $z_0 \in \mathbb{D}$ such that $h_b(z_0) = (0, -\sigma^2)$. Then

$$b(z_0) = -z_0 \quad \text{and} \quad z_0^2 = \sigma^2.$$

If $z_0 = \sigma$, then the conditions (i), (ii), and (iii) follow if we set $m = b$. Otherwise, if $z_0 = -\sigma$, the conclusions follow if we set $m = b^{-1}$. $\square$

Lemma 4.14 *If $\sigma > 0$ and $m \in \operatorname{Aut} \mathbb{D}$, then $m(\sigma) = -\sigma$ if and only if $m = b_\sigma \circ r_\tau \circ b_\sigma$ for some $\tau \in \mathbb{T}$. Furthermore, if $m(\sigma) = -\sigma$, then m is hyperbolic if and only if*

$$\left|\frac{\tau - 1}{\tau + 1}\right| < \frac{2\sigma}{1 - \sigma^2}, \tag{4.4}$$

and m is parabolic if and only if

$$\left|\frac{\tau - 1}{\tau + 1}\right| = \frac{2\sigma}{1 - \sigma^2}. \tag{4.5}$$

Proof

$$
\begin{aligned}
m(\sigma) = -\sigma \quad &\Leftrightarrow \quad m(b_\sigma^{-1}(0)) = b_\sigma(0) \\
&\Leftrightarrow \quad \left(b_\sigma^{-1} \circ m \circ b_\sigma^{-1}\right)(0) = 0 \\
&\Leftrightarrow \quad b_\sigma^{-1} \circ m \circ b_\sigma^{-1} = r_\tau \quad \text{for some } \tau \in \mathbb{T} \\
&\Leftrightarrow \quad m = b_\sigma \circ r_\tau \circ b_\sigma \quad \text{for some } \tau \in \mathbb{T},
\end{aligned}
$$

which proves the first assertion of the lemma.

To prove the second assertion, by direct calculation,

$$m(z) = \frac{\tau + \sigma^2}{1 + \tau\sigma^2} \frac{z - \frac{\sigma(\tau+1)}{\tau+\sigma^2}}{1 - \frac{\sigma(\tau+1)}{1+\tau\sigma^2} z}.$$

Consequently, by Lemma 4.1 (iii), *m* is hyperbolic if and only if

$$
\begin{aligned}
&\left|\frac{\tau + \sigma^2}{1 + \tau\sigma^2} - 1\right| < 2 \left|\frac{\sigma(\tau + 1)}{1 + \tau\sigma^2}\right| \\
\Leftrightarrow \quad &\left|\tau + \sigma^2 - (1 + \tau\sigma^2)\right| < 2 \left|\sigma(\tau + 1)\right| \\
\Leftrightarrow \quad &\left|\frac{\tau - 1}{\tau + 1}\right| < \frac{2\sigma}{1 - \sigma^2},
\end{aligned}
$$

that is, inequality (4.4) holds.

A similar calculation using Lemma 4.1 (ii) shows that *m* is parabolic if and only if Eq. (4.5) holds.   □

4.4 The Curve of Purely Balanced Directions

If τ_1 and τ_2 are in $\mathbb{T}$, we denote by (τ_1, τ_2) the open segment of points in $\mathbb{T}$ extending from τ_1 to τ_2 in the counterclockwise direction. In similar fashion we may define half open and closed segments in the circle. For $\sigma \in (0, 1)$, let τ_σ^+ denote the unique point in $\mathbb{T}$ that satisfies Eq. (4.5) and $\operatorname{Im} \tau_\sigma^+ > 0$ and let τ_σ^- denote the unique point in $\mathbb{T}$ satisfying Eq. (4.5) and $\operatorname{Im} \tau_\sigma^- < 0$. When $\sigma \in (0, 1)$ and $\tau \in \mathbb{T}$, we define $m_{\sigma,\tau}$ by

$$m_{\sigma,\tau} = b_\sigma \circ r_\tau \circ b_\sigma. \tag{4.6}$$

Lemma 4.15 *The map* $(\sigma, \tau) \mapsto B_{m_{\sigma,\tau}}$ *is injective from the set*

$$X \overset{\text{def}}{=} \{(\sigma, \tau) : 0 < \sigma < 1, \ \tau \in (\tau_\sigma^-, \tau_\sigma^+)\}$$

into the set of purely balanced geodesics in G.

Proof It follows from Lemmas 4.14 and 4.13 that, for $(\sigma, \tau) \in X$, the geodesic $B_{m_{\sigma,\tau}}$ is purely balanced.

Suppose that two points (σ, τ) and (σ', τ') in X give rise to the same purely balanced geodesic B. Notice firstly that B only meets the flat geodesic $F^0 = \{0\} \times \mathbb{D}$ in a single point, else, by the uniqueness of geodesics through any pair of distinct points, we have $B = F_0$, contrary to the hypothesis that B is purely balanced. By Lemma 4.13, both $(0, -\sigma^2)$ and $(0, -(\sigma')^2)$ lie in B, hence coincide. Since σ, σ' are positive, it follows that $\sigma = \sigma'$.

To see that $\tau = \tau'$, calculate a tangent vector to B at $(0, -\sigma^2)$. We find that

$$h'_{m_{\sigma,\tau}}(\sigma) = (1 + \tau, -\sigma(1 - \tau)).$$

It follows that, for some nonzero complex number c,

$$(1 + \tau, -\sigma(1 - \tau)) = c(1 + \tau', -\sigma(1 - \tau')),$$

from which it follows that $\tau = \tau'$. □

With the notations of the previous paragraph, the following proposition is an immediate consequence of Lemmas 4.13 and 4.14.

Proposition 4.16 *Let $\sigma \in (0, 1)$ and let D be a geodesic in G. Then D is a purely balanced geodesic passing through $(0, -\sigma^2)$ if and only if there exists $\tau \in (\tau_\sigma^-, \tau_\sigma^+)$ such that $D = B_{m_{\sigma,\tau}}$. Moreover D is an exceptional geodesic passing through $(0, -\sigma^2)$ if and only if $D = B_{m_{\sigma,\tau}}$ where either $\tau = \tau_\sigma^+$ or $\tau = \tau_\sigma^-$.*

We refer to geodesics of the form $B_{m_{\sigma,\tau}}$ as being in *standard position*. The following proposition states that any purely balanced or exceptional geodesic with a point prescribed in it can be moved to a geodesic in standard position by an automorphism of G.

Proposition 4.17 *Let D be a geodesic in G that is purely balanced (or exceptional) and let $\lambda \in D$. There exist $\sigma \in (0, 1)$, $\tau \in (\tau_\sigma^-, \tau_\sigma^+)$ (or $\tau \in \{\tau_\sigma^-, \tau_\sigma^+\}$, respectively), and $\gamma \in \operatorname{Aut} G$ such that $\gamma(\lambda) = (0, -\sigma^2)$ and $\gamma(D) = B_{m_{\sigma,\tau}}$.*

Proof Let $\lambda = \pi(z)$ where $z = (z_1, z_2) \in \mathbb{D}^2$. Since λ is in the purely balanced or exceptional geodesic D, $\lambda \notin \mathcal{R}$. Consequently, $z_1 \neq z_2$ and by the intermediate value theorem, we may choose $\sigma \in (0, 1)$ such that

$$d(\sigma, -\sigma) = d(z_1, z_2).$$

Choose $b \in \operatorname{Aut} \mathbb{D}$ such that

$$b(z_1) = \sigma \quad \text{and} \quad b(z_2) = -\sigma.$$

By Proposition 3.1, $\gamma_b \in \mathrm{Aut}\, G$. Furthermore, by formula (3.1),

$$\gamma_b(\lambda) = (b(z_1) + b(z_2), b(z_1)b(z_2)) = (0, -\sigma^2).$$

As B is assumed purely balanced or exceptional, so also is $\gamma_b(B)$, and $(0, -\sigma^2) \in \gamma_b(\mathbb{D})$. Therefore, the proposition follows by Proposition 4.16. □

Corollary 4.18 *If $\lambda \in \mathcal{R}$, then there are no purely balanced or exceptional directions at λ. If $\lambda \in G \setminus \mathcal{R}$, then there are exactly two exceptional directions at λ, and the purely balanced directions at λ form a simple smooth curve in CP^1 connecting the exceptional directions.*

Proof As no purely balanced or exceptional geodesic meets $\mathcal{R}$, there are no purely balanced or exceptional directions at points in $\mathcal{R}$. To prove the second assertion of the corollary, by Proposition 4.17, it suffices to prove the case when $\lambda = (0, -\sigma^2)$ for some $\sigma \in (0, 1)$. But this case is an immediate consequence of Proposition 4.16. □

5 Orthogonality in G

In this section, we shall study the geodesics that are orthogonal to a fixed flat geodesic.

5.1 Flat Geodesics

Recall that the flat geodesics of G were defined in Sect. 3.3 to be the geodesics D that have a unique royal point λ and furthermore are such that $\lambda \in \mathcal{R}$. It can be shown that flat geodesics are truly 'flat': they have the form[11]

$$F^\beta \stackrel{\text{def}}{=} \{(\beta + \bar{\beta}z, z) : z \in \mathbb{D}\} \tag{5.1}$$

for some $\beta \in \mathbb{D}$. Let us check that, for every $\beta \in \mathbb{D}$, this set is indeed a flat geodesic of G according to the definition. Firstly, if $(s, p) = (\beta + \bar{\beta}z, z)$ for some $z, \beta \in \mathbb{D}$, then $z = p$ and

$$s - \bar{s}p = \beta + \bar{\beta}p - (\bar{\beta} + \beta\bar{p}) = \beta(1 - |p|^2),$$

whence

$$|s - \bar{s}p| < 1 - |p|^2.$$

Hence, by the criterion (2.3), $(s, p) \in G$. Thus, $F^\beta \subset G$. Secondly, F^β is the range of the analytic disc $h(z) = (\beta + \bar{\beta}z, z)$ in G, and h has an obvious holomorphic left inverse, to wit, the second coordinate function. Thus, F^β is a geodesic of G.

---

[11] In earlier papers we *defined* the flat geodesics to be the F^β and showed that they are characterized by the property that they have a single royal point, which lies in $\mathcal{R}$. Here, we reverse the process in order to bring out the geometric nature of flatness.

It remains to show that F^β meets $\mathcal{R}^-$ at exactly one point λ, and that $\lambda \in \mathcal{R}$. Indeed, $(s, p) = h(z) \in \mathcal{R}^-$ if and only if

$$0 = s^2 - 4p = (\beta + \bar{\beta} z)^2 - 4z = \bar{\beta}^2 z^2 + 2(|\beta|^2 - 2)z + \beta^2.$$

If $\beta = 0$ then this equation for z has the unique root $z = 0$, so that F^0 meets $\mathcal{R}^-$ at the unique point $(0, 0)$. Otherwise, the quadratic equation has two roots, whose product has modulus one and whose sum has modulus greater than two. It follows that the equation has exactly one root in $\mathbb{D}$ and none in $\mathbb{T}$. Thus, in either case, F^β has a unique royal point, and that point lies in $\mathcal{R}$, which is to say that F^β is a flat geodesic.

Proposition 5.1 *Aut G permutes the set of flat geodesics of G. Indeed, for any $\beta \in \mathbb{D}$,*

$$\gamma_{b_\alpha}(F^\beta) = F^0,$$

where

$$\alpha = \frac{\beta}{1 + \sqrt{1 - |\beta|^2}} \tag{5.2}$$

is the hyperbolic midpoint of 0 and β in $\mathbb{D}$.

Proof By Corollary 3.2(3), every automorphism of G maps every flat geodesic to a flat geodesic.

Aut G acts transitively on $\mathcal{R}$, by reason of the equation

$$\gamma_m(2z, z^2) = (2m(z), m(z)^2)$$

for every $m \in$ Aut $\mathbb{D}$ and $z \in \mathbb{D}$. Hence, if $\lambda, \mu \in \mathcal{R}$, then we can find $m \in$ Aut $\mathbb{D}$ such that $\gamma_m(\lambda) = \mu$. It follows that $\gamma_m(F_\lambda)$ is a flat geodesic that contains $\mu \in \mathcal{R}$. Since there is a *unique* flat geodesic through any point of G, we infer that $\gamma_m(F_\lambda) = F_\mu$. Thus, Aut G permutes the flat geodesics of G.

It is not too arduous to calculate that, if $(s, p) \in F^\beta$ and α is given by Eq. (5.2), then

$$\gamma_{b_\alpha}(s, p) = \left(0, \frac{p - \alpha s + \alpha^2}{1 - \bar{\alpha} s + \bar{\alpha}^2 p}\right)$$

which is in F^0. $\qquad\square$

5.2 The Sharp Direction

Since every orbit of Aut G meets the flat geodesic F^0 and the sharp direction is covariant, the sharp direction at any point in G can be derived from the following statement.

Proposition 5.2 *For any* $\lambda \in F^0$,

$$\sharp_\lambda = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathbb{C}. \tag{5.3}$$

Proof Consider the point $\lambda = (0, p) \in F^0$. For any nonzero $\alpha \in \mathbb{C}$, let

$$f_\alpha(t) = \gamma_{b_{t\alpha}}(\lambda) \qquad \text{for } t \in \mathbb{R} \text{ such that } |t\alpha| < 1.$$

Then $f_\alpha(t)$ describes a smooth path in $\mathrm{Orb}(\lambda)$ such that $f_\alpha(0) = \lambda$, and hence $f_\alpha'(0) \in T_\lambda \, \mathrm{Orb}(\lambda)$.

Let $w \in \mathbb{D}$ satisfy $w^2 = -p$, so that $\lambda = \pi(-w, w)$. A short calculation shows that

$$\gamma_{b_\alpha}(\lambda) = \pi(b_\alpha(w), b_\alpha(-w)) = \frac{1}{1 + \bar{\alpha}^2 p} \left(-2(\alpha + \bar{\alpha}p), p + \alpha^2 \right).$$

Replace α by $t\alpha$, differentiate with respect to t and set $t = 0$ to obtain

$$f_\alpha'(0) = (-2(\alpha + \bar{\alpha}p), 0) \quad \in T_\lambda \, \mathrm{Orb}(\lambda)$$

for every $\alpha \in \mathbb{C} \setminus \{0\}$. By suitable choices of α, we deduce that both $(1, 0)$ and $(i, 0)$ lie in $T_\lambda \, \mathrm{Orb}(\lambda)$. Thus, the complex linear span of $(1, 0)$ is contained in $T_\lambda \, \mathrm{Orb}(\lambda)$, and statement (5.3) follows. $\qquad \square$

Corollary 5.3 *For every* $\lambda \in G$, $\sharp_\lambda \neq \flat_\lambda$.

Proof It is enough to prove it for every $\lambda \in F^0$. For such λ,

$$\sharp_\lambda = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathbb{C}, \quad \flat_\lambda = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mathbb{C}. \qquad \square$$

It follows that the tangent bundle TG is the direct sum of the sharp bundle $\lambda \mapsto \sharp_\lambda$ and the flat bundle $\lambda \mapsto \flat_\lambda$.

We remark that a formula for $\sharp_\lambda$ in general is given in [5, Proposition 1.16]. If $\lambda = (s, p) \in F^\beta$, then

$$\sharp_\lambda = \begin{pmatrix} 1 \\ \dfrac{\beta - \frac{1}{2}s}{1 - \frac{1}{2}\bar{\beta}s} \end{pmatrix} \mathbb{C}.$$

5.3 Orthogonality and Irrotational Automorphisms

Recall that a geodesic D is said to be orthogonal to a flat geodesic F if D meets F at a point λ and $T_\lambda D = \sharp_\lambda$. In view of Proposition 5.1, in studying any flat geodesic F,

one may often reduce to the case that

$$F = F^0 = \{(0, p) : p \in \mathbb{D}\},$$

the flat geodesic passing through the origin.

Proposition 5.4 *A geodesic D is orthogonal to F^0 if and only if $D = B_m$ for some irrotational $m \in \operatorname{Aut} \mathbb{D}$.*

Proof First assume that $D \perp F^0$. Choose $\lambda \in D \cap F^0$. Let $\lambda = (0, p_0)$ and choose $\alpha \in \mathbb{D}$ such that $\alpha^2 = -p_0$. Let $m \in \operatorname{Aut} \mathbb{D}$ where $-m$ is hyperbolic reflection about α. Observe that Lemma 4.6 implies that m is irrotational.

To see that $D = B_m$ note that Condition (ii) in Lemma 4.6 implies that $m(\alpha) = -\alpha$, which implies that $h_m(\alpha) = \lambda$. Therefore, $\lambda \in B_m$. Also, Condition (ii) in Lemma 4.6 implies that $m'(\alpha) = 1$. Therefore,

$$h'_m(\alpha) = (1 + m'(\alpha), m(\alpha) + \alpha m'(\alpha)) = (1 + 1, -\alpha + \alpha) = (2, 0),$$

which implies that $T_\lambda B_m = \sharp_\lambda = T_\lambda D$. As D and B_m both pass through the point λ and have the same tangent space at λ, it follows from Theorem 3.3 that $D = B_m$.

Now assume that m is irrotational. We need to show that $B_m \perp F^0$. Condition (ii) in Lemma 4.6 implies that there exists $\alpha \in \mathbb{D}$ such that $m(\alpha) = -\alpha$. As,

$$h_m(\alpha) = (0, -\alpha^2) \in F^0,$$

it follows that B_m meets F^0 at the point $\lambda = (0, -\alpha^2)$. In addition, Condition (ii) in Lemma 4.6 implies that $m'(\alpha) = 1$. Therefore, calculating as before, we have

$$h'_m(\alpha) = (2, 0),$$

which implies that $T_\lambda B_m = \sharp_\lambda$. As B_m meets F^0 at the point λ and $T_\lambda B_m = \sharp_\lambda$, it follows that B_m is orthogonal to F^0. □

Corollary 5.5 *If $\mu \in \mathcal{R}$, then $\sharp_\mu = T_\mu \mathcal{R}$. If $\mu \in G \setminus \mathcal{R}$, then the geodesic through μ with direction $\sharp_\mu$ is purely balanced.*

Proof Fix $\mu \in G$, let F be the flat geodesic passing through μ, and let D be the geodesic such that $\mu \in D$ and $T_\mu D = \sharp_\mu$. We wish to show that $D = \mathcal{R}$ when $\mu \in \mathcal{R}$, and otherwise, that D is purely balanced.

Choose $\gamma \in \operatorname{Aut} G$ such that $\lambda = \gamma(\mu) \in F^0$. As γ is an automorphism and $D \perp F$, $\gamma(D) \perp F^0$. Therefore, by Proposition 5.4, there exists an irrotational m such that $\gamma(D) = B_m$.

As m is irrotational, by Proposition 4.7, there are only two possibilities:

(i) $m = \operatorname{id}_\mathbb{D}$ and $B_m = \mathcal{R}$, and
(ii) m is hyperbolic and B_m is purely balanced.

In case (i), since γ fixes $\mathcal{R}$, $\mu \in \mathcal{R}$ and $\sharp_\mu = T_\mu \mathcal{R}$. In case (ii), $\mu \notin \mathcal{R}$, and $\sharp_\mu$ is purely balanced. □

The following result is essentially Corollary 5.5 stated in different language.

Corollary 5.6 *The royal geodesic is orthogonal to every flat geodesic. If a geodesic D is orthogonal to a flat geodesic, then either $D = \mathcal{R}$ or D is purely balanced.*

The following result implies that at points not on the royal geodesic, the sharp direction is the midpoint of the curve of purely balanced directions.

Corollary 5.7 *For the curve of purely balanced geodesics $B_\tau = B_{m_{\sigma,\tau}}$ constructed in Proposition 4.16, $T_{(0,-\sigma^2)}B_\tau = \sharp_{(0,-\sigma^2)}$ if and only if $\tau = 1$.*

Proof Proposition 5.4 implies that $T_{(0,-\sigma^2)}B_\tau = \sharp_{(0,-\sigma^2)}$ if and only if $m_{\sigma,\tau}$ is irrotational. But Lemma 4.6 implies that $m_{\sigma,\tau}$ is irrotational if and only if $\tau = m'_{\sigma,\tau}(\sigma) = 1$. □

5.4 Foliation of G by Orthogonal Geodesics

If F is a geodesic in $\mathbb{R}^2$ (that is, a line), then $\mathbb{R}^2$ is foliated by the geodesics orthogonal to F. Theorem 5.10 is an analogous result in G.

Lemma 5.8 *If F is a flat geodesic in G, and D_1 and D_2 are geodesics that are orthogonal to F, then either $D_1 = D_2$ or $D_1 \cap D_2 = \varnothing$.*

Proof As Aut G acts transitively on flat geodesics, we may assume without loss of generality that $F = F^0$. Let D_1 and D_2 be geodesics that are orthogonal to F. By Proposition 5.4, there exist irrotational automorphisms m_1 and m_2 of $\mathbb{D}$ such that $D_1 = B_{m_1}$ and $D_2 = B_{m_2}$.

If $B_{m_1} \cap B_{m_2} \neq \varnothing$, then there exist z_1 and z_2 in $\mathbb{D}$ such that $h_{m_1}(z_1) = h_{m_2}(z_2)$, equations which imply either

$$z_1 = z_2 \quad \text{and} \quad m_1(z_1) = m_2(z_2) \tag{5.4}$$

or

$$z_1 = m_2(z_2) \quad \text{and} \quad m_1(z_1) = z_2. \tag{5.5}$$

If Eq. (5.4) holds then $m_1(z_1) = m_2(z_1)$, or equivalently, $(m_2^{-1} \circ m_1)(z_1) = z_1$. Hence, $m_2^{-1} \circ m_1$ is elliptic, and Remark 4.4(1) implies that $m_1 = m_2$. Thus, $D_1 = D_2$.

If Eq. (5.5) holds, then $(m_2 \circ m_1)(z_1) = z_1$ and we see that $m_2 \circ m_1$ is elliptic. Then Remark 4.4(1) implies that $m_1 = m_2^{-1}$, and again, $D_1 = D_2$. □

Lemma 5.9 *Fix a flat geodesic F and a point $\mu \in G \backslash F$. There exists a purely balanced geodesic D such that D is orthogonal to F and $\mu \in D$.*

Proof Without loss of generality, we may assume that $F = F^0$. Choose $z_1, z_2 \in \mathbb{D}$ such that

$$\mu = \pi(z_1, z_2) = (z_1 + z_2, z_1 z_2).$$

Represent $\pi(z_1, -z_2)$ in flat coordinates,

$$\pi(z_1, -z_2) = (z_1 - z_2, -z_1 z_2) = (\beta + \bar{\beta} p_0, p_0),$$

where $\beta, p_0 \in \mathbb{D}$. We have

$$z_1 - z_2 = \beta - \bar{\beta} z_1 z_2,$$

which implies that

$$z_2 = \frac{z_1 - \beta}{1 - \bar{\beta} z_1} = b_\beta(z_1).$$

If we set $D = B_{b_\beta}$, then, as $h_{b_\beta}(z_1) = \mu$, we can assert that $\mu \in D$. Furthermore, as b_β is irrotational, Proposition 5.4 guarantees that D is orthogonal to F. $\qquad\square$

By combining the previous two lemmas we obtain the following theorem.

Theorem 5.10 *If F is a flat geodesic, then G is foliated by the geodesics in G that are orthogonal to F.*

6 Distinguished Geodesics in _G_

In the previous section, we studied the purely balanced geodesics that are orthogonal to a fixed flat geodesic. In this section, in contrast, we fix a purely balanced geodesic D and study the flat geodesics F such that D is orthogonal to F. This leads to the remarkable discovery that such flat geodesics are naturally parametrized by a real geodesic in D.

Definition 6.1 For any geodesic D, we say that a point $\lambda \in D$ is a *sharp point in D* if D is orthogonal to the flat geodesic passing through λ. The set of sharp points in D is denoted by $\sharp(D)$.

Thus, $\sharp(D) = \{\lambda \in D : T_\lambda D = \sharp_\lambda\}$. Corollary 5.5 implies that $\sharp(\mathcal{R}) = \mathcal{R}$ and that, if $D \neq \mathcal{R}$, then $\sharp(D)$ is non-empty only if D is purely balanced. We therefore restrict attention to the case that D is a purely balanced geodesic B_m.

Proposition 6.3, which gives a description of the sharp points in a purely balanced geodesic B, represents the third promised relationship between $B = B_m$ and the fixed points of m.

If B is a purely balanced geodesic, we may define a real geodesic in B in the following manner. By Proposition 4.7, there exists hyperbolic $m \in \operatorname{Aut} \mathbb{D}$ such that $B = B_m$. Let $\eta = \{\eta_1, \eta_2\}$ denote the set of fixed points of m and let C_η denote the real

hyperbolic geodesic in $\mathbb{D}$ that has η_1 and η_2 as endpoints. Finally, we define $\Xi_B \subset B$ by

$$\Xi_B = h_m(C_\eta),$$

where, for $z \in \mathbb{D}$, $h_m(z) = (z + m(z), zm(z))$.

Lemma 6.2 *If B is a purely balanced geodesic and $B = B_m$, then the definition of Ξ_B does not depend on the choice of m.*

Proof Suppose that $m, q \in \text{Aut } \mathbb{D}$ are such that $B_m = B = B_q$. By Lemma 4.8, either $q = m$ or $q = m^{-1}$. In the former case it is immediate that m and q yield the same curve Ξ_B. In the latter case, since the fixed points ζ of m^{-1} coincide with those η of m, we have $\eta = \zeta$, and so $C_\eta = C_\zeta$. Moreover, since m is a hyperbolic isometry of $\mathbb{D}$, m maps geodesics to geodesics in $\mathbb{D}$, and so $m(C_\eta) = C_\eta = C_\zeta$. For any $z \in \mathbb{D}$, we have $h_{m^{-1}}(m(z)) = h_m(z)$, and therefore

$$h_q(C_\zeta) = h_{m^{-1}}(m(C_\eta)) = h_m(C_\eta) = \Xi_B.$$

$\square$

Proposition 6.3 *If B is a purely balanced geodesic, then $\sharp(B) = \Xi_B$.*

Proof We first show that $\sharp(B) \subseteq \Xi_B$. Fix $\lambda \in \sharp(B)$. By Proposition 4.17, we may assume that B is in standard position, $\lambda = (0, -\sigma^2)$ and $B = B_{m_{\sigma,\tau}}$ for some $\sigma \in (0, 1)$ and $\tau \in \mathbb{T}$. Furthermore, as $\lambda \in \sharp(B)$, Corollary 5.7 implies that $\tau = 1$. Since the fixed points of $m_{\sigma,1}$ are ± 1, $C_\eta = (-1, 1)$. Therefore, as $\sigma \in (-1, 1)$,

$$\lambda = (0, -\sigma^2) = h_{m_{\sigma,1}}(\sigma) \in h_{m_{\sigma,1}}(C_\eta) = \Xi_B.$$

We now turn to the proof that $\Xi_B \subseteq \sharp(B)$. Fix a purely balanced geodesic B with royal points $(2\xi, \xi^2)$ and $(2\eta, \eta^2)$. Assume that B is parametrized as in Eq. (4.3), so that $B = B_m$ for some hyperbolic $m \in \text{Aut } \mathbb{D}$, with fixed points at ξ and η in $\mathbb{T}$.
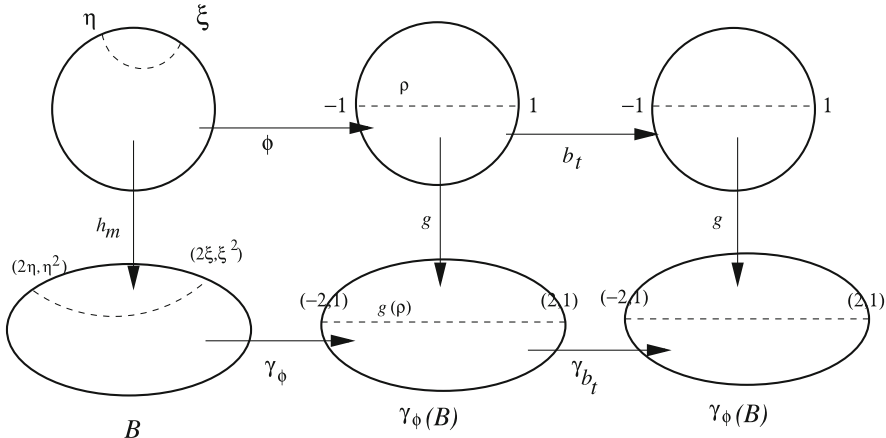
Choose $\varphi \in \operatorname{Aut} \mathbb{D}$ satisfying $\varphi(\xi) = 1$ and $\varphi(\eta) = -1$. Since $\varphi \circ m \circ \varphi^{-1}(1) = 1$ and $\varphi \circ m \circ \varphi^{-1}(-1) = -1$, it follows that there exists $r \in (-1, 1)$ such that $\varphi \circ m \circ \varphi^{-1} = b_r$. Furthermore, $\gamma_\varphi(B)$ is parametrized by the function g defined by

$$g(z) = (z + b_r(z), zb_r(z)), \qquad z \in \mathbb{D}. \tag{6.1}$$

Since $b_r(-1) = -1$, $b_r(1) = 1$, and $b_r|[-1, 1]$ is strictly increasing, it follows from the Intermediate Value Theorem that there exists a unique $\rho \in (-1, 1)$ such that $b_r(\rho) = -\rho$. By Lemma 4.6, $b_r'(\rho) = 1$, and we deduce from Eq. (6.1) that

$$g(\rho) = (0, -\rho^2) \quad \text{and} \quad g'(\rho) = (2, 0).$$

Since g parametrizes $\gamma_\varphi(B)$, $g'(\rho) \in T_{g(\rho)}\gamma_\varphi(B)$. Therefore, by Eq. (5.3), $g(\rho)$ is a sharp point in $\gamma_\varphi(B)$.

Now fix $t \in (-1, 1)$. Since $b_t \circ b_r = b_r \circ b_t$,

$$\begin{aligned}
\gamma_{b_t} \circ g(z) &= \gamma_{b_t}(z + b_r(z), zb_r(z)) \\
&= (b_t(z) + b_t \circ b_r(z), b_t(z)b_t \circ b_r(z)) \\
&= (b_t(z) + b_r \circ b_t(z), b_t(z)b_r \circ b_t(z)) = g \circ b_t(z),
\end{aligned}$$

for all $z \in \mathbb{D}$. Hence $\gamma_{b_t} \circ g(\rho) = g \circ b_t(\rho)$, and the tangent space

$$T_{g \circ b_t(\rho)}\gamma_\varphi(B) = \gamma_{b_t}' \circ g(\rho)\, T_{g(\rho)}\gamma_\varphi(B).$$

Therefore, as $g(\rho)$ is a sharp point in $\gamma \circ \varphi(B)$,

$$\begin{aligned}
T_{g \circ b_t(\rho)}\gamma \circ \varphi(B) &= \gamma_{b_t}' \circ g(\rho)\, T_{g(\rho)}\gamma_\varphi(B) \\
&= \gamma_{b_t}' \circ g(\rho)\, \sharp_{g(\rho)} = \sharp_{g \circ b_t(\rho)},
\end{aligned}$$

that is,

$$g \circ b_t(\rho) \in \sharp(\gamma_\varphi(B)). \tag{6.2}$$

Since Eq. (6.2) holds for all $t \in (-1, 1)$ and

$$\{b_t(\rho) : t \in (-1, 1)\} = (-1, 1) = C(\{-1, 1\}),$$

it follows that $g((-1, 1)) \subseteq \sharp(\gamma_\varphi(B))$, which is to say that

$$\Xi_{\gamma_\varphi(B)} \subseteq \sharp(\gamma_\varphi(B)).$$

As γ_φ is an automorphism of G, $\Xi_B \subseteq \sharp(B)$, as was to be proved. $\qquad \square$

Definition 6.4 A *simple real geodesic* in G is a curve of the form $h(C)$ where $h : \mathbb{D} \to G$ is a hyperbolic isometry and C is a real geodesic in $\mathbb{D}$.

A *distinguished geodesic in* G is a simple real geodesic in G whose endpoints lie in the edge E of the distinguished boundary of G.

Thus, a simple real geodesic in G is a curve C in G such that, for any pair of points $\lambda, \mu \in C$, the segment of C joining λ and μ achieves the minimum hyperbolic length of any curve in G joining λ and μ. Note that not all real geodesics are simple.

Two types of distinguished geodesic Ξ in G are as follows. We say Ξ is a *royal distinguished* geodesic if $\Xi = h(C)$ where $h : \mathbb{D} \to \mathcal{R}$ is a hyperbolic isometry and C is a real hyperbolic geodesic in $\mathbb{D}$. We say Ξ is a *purely balanced distinguished geodesic* if $\Xi = \Xi_B$ for some purely balanced geodesic B.

Proposition 6.5 Ξ *is a distinguished geodesic in* G *if and only if it is either a royal or a purely balanced distinguished geodesic.*

Lemma 6.6 *If* $\lambda \in G \setminus \mathcal{R}$*, then there exists a purely balanced geodesic* B *such that* $\lambda \in \Xi_B$.

Proof Let B be the geodesic in G such that $\lambda \in B$ and $T B_\lambda = \sharp_\lambda$. By Corollary 5.5, B is purely balanced. Also, since $\lambda \in \sharp(B)$, Proposition 6.3 implies that $\lambda \in \Xi_B$. □

Lemma 6.7 *If* B_1 *and* B_2 *are distinct purely balanced geodesics in* G*, then* $\Xi_{B_1} \cap \Xi_{B_2} = \varnothing$.

Proof If $\lambda \in \Xi_{B_1} \cap \Xi_{B_2}$, then

$$T_\lambda B_1 = \sharp_\lambda = T_\lambda B_2.$$

Hence, Theorem 3.3 implies that $B_1 = B_2$. □

The above two lemmas have the following immediate consequence.

Theorem 6.8 *The purely balanced distinguished geodesics foliate* $G \setminus \mathcal{R}$.

As a consequence of Theorem 6.8, we may seek coordinates for $G \setminus \mathcal{R}$ based on a parametrization of the purely balanced distinguished geodesics in G. Let

$$X = \{\{\xi_1, \xi_2\} : \xi_1, \xi_2 \in E, \xi_1 \neq \xi_2\}$$

Since E is a circle, topologically, X is homeomorphic to $M \setminus E$, a Möbius band without boundary. For each point $\xi = \{\xi_1, \xi_2\} \in X$, we may construct a smooth parametrization

$$t \mapsto B_{\xi, t}, \qquad t \in (0, 1),$$

of the family of purely balanced geodesics B such that $B \cap E = \xi$. Finally, for each $\xi \in X$ and each $t \in (0, 1)$, we may smoothly parametrize $\Xi_{\xi, t}$ using $(0, 1)$. These parametrizations lead to the following result.

Theorem 6.9 $G \setminus \mathcal{R}$ *is naturally homeomorphic to* $X \times (0, 1) \times (0, 1)$ *via the parametrizations described above.*

7 The Closest Point Property in *G*

In [6, Lemma 9.8], it was shown that if F is a flat geodesic in G and $\mu \in G \setminus F$, then

$$\inf_{\lambda \in F} d_G(\lambda, \mu) \text{ is attained at a point } \lambda_0 \in F. \tag{7.1}$$

Further[12], it was shown that if λ_0 and μ satisfy Eq. (7.1), and D denotes the geodesic passing through μ and λ_0, then

$$\mu \in \mathcal{R} \Rightarrow D = \mathcal{R} \quad \text{and} \quad \mu \notin \mathcal{R} \Rightarrow D \text{ is purely balanced.} \tag{7.2}$$

This result prompts the following definition.

Definition 7.1 Let F be a flat geodesic in G and let $\mu \in G$. We say that λ_0 is *a closest point in F to μ* if

$$d_G(\lambda_0, \mu) = \inf_{\lambda \in F} d_G(\lambda, \mu).$$

If D is a geodesic in G, we say that D *has the closest point property with respect to F* if F meets D in a point λ_0 and for some $\mu \in D \setminus \{\lambda_0\}$,

$$\lambda_0 \text{ is a closest point in } F \text{ to } \mu. \tag{7.3}$$

Theorem 7.9 implies that if D has the closest point property with respect to F, then, in fact, statement (7.3) holds for every $\mu \in D$.

7.1 Critical Pairs

Definition 7.2 Let $\lambda_0, \mu \in G$. We say that the pair (λ_0, μ) is a *critical pair* if for every admissible direction $v \in \flat_{\lambda_0}$,

$$\frac{d}{dt} d_G(\lambda_0 + tv, \mu) \Big|_{t=0} \geq 0. \tag{7.4}$$

Lemma 7.3 *Let F be a flat geodesic in G and let $\mu \in G \setminus F$. If λ_0 is a closest point in F to μ, then (λ_0, μ) is a critical pair.*

If $\lambda_0, \mu \in G$ and the geodesic D passing through λ_0 and μ is purely balanced, then we may let $D = B_m$, where m is hyperbolic with fixed points at η_1 and η_2. It then follows, by Lemma 4.12, that

$$d_G(\lambda_0, \mu) = \max_{\omega \in \{\bar{\eta}_1, \bar{\eta}_2\}} d(\Phi_\omega(\lambda_0), \Phi_\omega(\mu)).$$

---

[12] Corollary 7.10 asserts that λ_0 is unique.

Consequently, if $v \in \flat_{\lambda_0}$, the following formula holds:

$$\frac{d}{dt} d_G(\lambda_0 + tv, \mu)\Big|_{t=0} = \max_{\omega \in \{\bar{\eta}_1, \bar{\eta}_2\}} \frac{d}{dt} d(\Phi_\omega(\lambda_0 + tv), \Phi_\omega(\mu))\Big|_{t=0}. \quad (7.5)$$

Lemma 7.4 *Assume that* $\lambda_0 = (0, p_0)$ *is a point in* F_0, *the flat geodesic passing through the origin, and let* $\mu \in G \setminus F_0$. *Assume that the geodesic* D *passing through* λ_0 *and* μ *is purely balanced and let* $D = B_m$ *where m is hyperbolic with fixed points* η_1 *and* η_2. (λ_0, μ) *is a critical pair if and only if, for all* $\xi \in \mathbb{C}$,

$$either \quad \frac{d}{dt} d(p_0 + t\xi, w_1)\Big|_{t=0} \geq 0 \quad or \quad \frac{d}{dt} d(p_0 + t\xi, w_2)\Big|_{t=0} \geq 0, \quad (7.6)$$

where $w_1, w_2 \in \mathbb{D}$ *are defined by*

$$w_1 = \eta_1 \Phi_{\bar{\eta}_1}(\mu) \quad and \quad w_2 = \eta_2 \Phi_{\bar{\eta}_2}(\mu). \quad (7.7)$$

Proof First notice that $v \in \flat_{(\lambda_0, \mu)}$ if and only if $v = (0, d)$ for some $d \in \mathbb{C}$. Consequently, using Definition 7.2 and (7.5), we see that (λ_0, μ) is a critical pair if and only if for all $\xi \in \mathbb{C}$ there exists $\omega \in \{\bar{\eta}_1, \bar{\eta}_2\}$ such that

$$\frac{d}{dt} d(\Phi_\omega(0, p_0 + t\xi), \Phi_\omega(\mu))\Big|_{t=0} \geq 0. \quad (7.8)$$

Note that, if $\omega \in \mathbb{T}$,

$$d(\Phi_\omega(0, p_0 + t\xi), \Phi_\omega(\mu)) = d(\omega(p_0 + t\xi), \Phi_\omega(\mu))$$
$$= d(p_0 + t\xi, \bar{\omega}\Phi_\omega(\mu)).$$

Hence statement (7.8) becomes

for all $\xi \in \mathbb{C}$ there exists $\omega \in \{\bar{\eta}_1, \bar{\eta}_2\}$ such that $\dfrac{d}{dt} d(p_0 + t\xi, \bar{\omega}\Phi_\omega(\mu))\Big|_{t=0} \geq 0$,

which is equivalent to statement (7.6). □

A moment's thought reveals that statement (7.8) is equivalent to the assertion that w_2 is the hyperbolic reflection of w_1 about the point p_0. Thus, we obtain the following result.

Lemma 7.5 *Assume that* $\lambda_0 = (0, p_0)$ *is a point in the flat geodesic* F_0 *passing through the origin, and let* $\mu \in G \setminus F_0$. *Assume that the geodesic* D *passing through* λ_0 *and* μ *is purely balanced and let* $D = B_m$ *where m is hyperbolic with fixed points* η_1 *and* η_2. *Let* $w_1 = \eta_1 \Phi_{\bar{\eta}_1}(\mu)$ *and* $w_2 = \eta_2 \Phi_{\bar{\eta}_2}(\mu)$. *Then* (λ_0, μ) *is a critical pair if and only if* w_2 *is the hyperbolic reflection of* w_1 *about* p_0.

7.2 Critical Pairs and Orthogonality

Lemma 7.6 *Let $p_0 \in \mathbb{D} \setminus \{0\}$. A geodesic D is purely balanced and meets F_0 at the point $(0, p_0)$ if and only if there exists $\alpha \in \mathbb{D} \setminus \{0\}$ and hyperbolic $m \in \operatorname{Aut} \mathbb{D}$ such that $D = B_m$, $m(\alpha) = -\alpha$, and $\alpha^2 = -p_0$.*

Proof By Proposition 4.7, m is hyperbolic if and only if $D = B_m$ is purely balanced. In addition, $m(\alpha) = -\alpha$ and $\alpha^2 = -p_0$ if and only if $h_m(\alpha) = (0, p_0)$. □

Lemma 7.7 *Let $m \in \operatorname{Aut} \mathbb{D}$ be hyperbolic and assume that $\alpha \in \mathbb{D} \setminus \{0\}$ where $m(\alpha) = -\alpha$, and $\alpha^2 = -p_0$. There exists $\tau \in \mathbb{T}$ such that[13] $m = b_\alpha \circ r_\tau \circ b_\alpha$. Furthermore, if we let η_1, η_2 denote the fixed points of m and define*

$$\sigma = \eta_1 + \eta_2 \quad and \quad \pi = \eta_1 \eta_2,$$

then

$$\sigma = \frac{1 - |\alpha|^2}{\bar{\alpha}} \frac{1 - \tau}{1 + \tau} \quad and \quad \pi = -\frac{\alpha}{\bar{\alpha}}. \tag{7.9}$$

Proof That the general $m \in \operatorname{Aut} \mathbb{D}$ satisfying $m(\alpha) = -\alpha$ has the form $m = b_\alpha \circ r_\tau \circ b_\alpha$ follows by Schur reduction.

To prove Eq. (7.9), it suffices to show that the equation $m(x) = x$ is equivalent to the equation $x^2 - \sigma x + \pi = 0$.

$$
\begin{aligned}
m(x) = x \quad &\Leftrightarrow \quad (b_\alpha \circ r_\tau \circ b_\alpha)(x) = x \\
&\Leftrightarrow \quad \tau b_\alpha(x) = b_{-\alpha}(x) \\
&\Leftrightarrow \quad \tau(x - \alpha)(1 + \bar{\alpha}x) = (x + \alpha)(1 - \bar{\alpha}x) \\
&\Leftrightarrow \quad \bar{\alpha}(1 + \tau)x^2 - (1 - |\alpha|^2)(1 - \tau)x - \alpha(1 + \tau) = 0 \\
&\Leftrightarrow \quad x^2 - \frac{1 - |\alpha|^2}{\bar{\alpha}} \frac{1 - \tau}{1 + \tau}x - \frac{\alpha}{\bar{\alpha}} = 0 \\
&\Leftrightarrow \quad x^2 - \sigma x + \pi = 0.
\end{aligned}
$$

□

We remark that Eq. (7.9) implies that $\tau = 1$ if and only if $\sigma = 0$, a fact which also follows from Lemma 4.5 and Condition (ii) in Lemma 4.6. Also observe that the formulas (7.9) imply the formulas

$$\frac{1 + \tau}{1 - \tau}\sigma = \frac{1 - |\alpha|^2}{\bar{\alpha}} \quad and, since \ \tau \in \mathbb{T}, \quad \frac{1 + \tau}{1 - \tau}\bar{\sigma} = -\frac{1 - |\alpha|^2}{\alpha}. \tag{7.10}$$

---

[13] By the Chain Rule, $\tau = m'(\alpha)$.

Lemma 7.8 *Assume that* $\lambda_0 = (0, p_0)$ *is a point in* F_0, *the flat geodesic passing through the origin, and let* $\mu \in G \setminus F_0$. *Assume that the geodesic D passing through* λ_0 *and* μ *is purely balanced and let* $D = B_m$ *where m is hyperbolic with fixed points* η_1 *and* η_2. *If* (λ_0, μ) *is a critical pair, then D is orthogonal to* F_0.

Proof By Lemma 7.6, there exists $\alpha \in \mathbb{D}$ such that

$$m(\alpha) = -\alpha, \quad \text{and} \quad \alpha^2 = p_0, \tag{7.11}$$

equations that imply

$$h_m(\alpha) = (0, p_0). \tag{7.12}$$

As also $\mu \in B_m$, there exists $\beta \in \mathbb{D}$ such that

$$h_m(\beta) = \mu. \tag{7.13}$$

Lemma 4.12 guarantees that if we define m_1 and m_2 by

$$m_1 = (\eta_1 \Phi_{\bar{\eta}_1}) \circ h_m \quad \text{and} \quad m_1 = (\eta_2 \Phi_{\bar{\eta}_2}) \circ h_m \tag{7.14}$$

then $m_1, m_2 \in \text{Aut}\,\mathbb{D}$. Observe that

$$m_1(\alpha) = \eta_1 \Phi_{\bar{\eta}_1}(0, p_0) = p_0,$$

and using Eq. (7.7), that

$$m_1(\beta) = \eta_1 \Phi_{\bar{\eta}_1}(\mu) = w_1.$$

Similarly,

$$m_2(\alpha) = p_0 \quad \text{and} \quad m_2(\beta) = w_2.$$

Consequently, if we set $\varphi = m_2 \circ m_1^{-1}$,

$$\varphi(p_0) = p_0 \quad \text{and} \quad \varphi(w_1) = w_2.$$

As Lemma 7.5 guarantees that w_2 is the hyperbolic reflection of w_1 about p_0, it follows that φ is hyperbolic reflection about p_0. Therefore, as $\varphi'(p_0) = -1$,

$$m_1'(\alpha) + m_2'(\alpha) = 0 \tag{7.15}$$

Noting that, for $\eta \in \mathbb{T}$,

$$\nabla(\eta \Phi_{\bar{\eta}})(s, p) = \left(\frac{-2\eta + 2\bar{\eta}p}{(2 - \bar{\eta}s)^2}, \frac{2}{2 - \bar{\eta}s} \right),$$

we see that

$$
\begin{aligned}
m_1'(\alpha) &= \nabla(\eta_1 \Phi_{\bar{\eta}_1})(0, p_0) \cdot h_m'(\alpha) \\
&= \left(-\tfrac{1}{2}\eta_1 + \tfrac{1}{2}\bar{\eta}_1 p_0, 1 \right) \cdot \left(1 + m'(\alpha), \alpha(m'(\alpha) - 1) \right) \\
&= (1 + m'(\alpha))(-\tfrac{1}{2}\eta_1 + \tfrac{1}{2}\bar{\eta}_1 p_0) + \alpha(m'(\alpha) - 1).
\end{aligned}
$$

Likewise

$$
m_2'(\alpha) = (1 + m'(\alpha))(-\tfrac{1}{2}\eta_2 + \tfrac{1}{2}\bar{\eta}_2 p_0) + \alpha(m'(\alpha) - 1).
$$

Therefore, by Eq. (7.15) and Lemma 7.7,

$$
\begin{aligned}
0 &= m_1'(\alpha) + m_2'(\alpha) \\
&= (1 + m'(\alpha))(-\tfrac{1}{2}\eta_1 + \tfrac{1}{2}\bar{\eta}_1 p_0) + \alpha(m'(\alpha) - 1) \\
&\quad + (1 + m'(\alpha))(-\tfrac{1}{2}\eta_2 + \tfrac{1}{2}\bar{\eta}_2 p_0) + \alpha(m'(\alpha) - 1) \\
&= (1 + \tau)(-\tfrac{1}{2}\sigma - \tfrac{1}{2}\bar{\sigma}\alpha^2) + 2\alpha(\tau - 1) \qquad \text{by Lemma7.7} \\
&= -\frac{1 - \tau}{2}\left(\frac{1 + \tau}{1 - \tau}\sigma + \frac{1 + \tau}{1 - \tau}\bar{\sigma}\alpha^2 + 4\alpha \right) \\
&= -\frac{1 - \tau}{2}\left(\frac{1 - |\alpha|^2}{\bar{\alpha}} - \frac{1 - |\alpha|^2}{\alpha}\alpha^2 + 4\alpha \right) \quad \text{by Eq.(7.10)} \\
&= -\frac{1 - \tau}{2\bar{\alpha}}\left(1 - |\alpha|^2 \right)^2.
\end{aligned}
$$

Consequently, as $m'(\alpha) = \tau$, $m'(\alpha) = 1$ and Lemma 4.6 implies that m is irrotational. By Proposition 5.4, D is orthogonal to F_0. □

7.3 The Closest Point Property and Orthogonality

Theorem 7.9 *Let F be a flat geodesic in G. Assume that D is a geodesic in G that meets F at the point λ_0. The following are equivalent.*

(i) *For every $\mu \in D$, λ_0 is a closest point in F to μ;*
(ii) *there exists $\mu \in D$ such that λ_0 is a closest point in F to μ;*
(iii) *D is orthogonal to F.*

Proof Clearly, (i) implies (ii). If (ii) holds, then (7.2) implies that D is purely balanced and Lemma 7.3 implies that (λ_0, μ) is a critical pair. Therefore, Lemma 7.8 implies that (iii) holds.

Now suppose that (iii) holds. Fix $\mu \in \mathbb{D}$. Choose $\lambda_1 \in F$ such that λ_1 is a closest point in F to μ. By (ii) implies (iii), if D_1 is the geodesic passing through λ_1 and μ, then D_1 is orthogonal to F. As D and D_1 are both orthogonal to F and $\mu \in D \cap D_1$, Lemma 5.8 implies that $\lambda_1 = \lambda_0$. Therefore, λ_0 is a closest point in F to μ. □

Corollary 7.10 *Assume that F is a flat geodesic in G and $\mu \in G$. If λ_1 and λ_2 are both a closest point in F to μ, then $\lambda_1 = \lambda_2$.*

Proof Suppose

$$\inf_{\lambda \in F} d_G(\lambda, \mu)$$

is attained at two points λ_1 and λ_2 in F. If we let D_1 be the geodesic passing through λ_1 and μ and let D_2 be the geodesic passing through λ_2 and μ, then the equivalence of Conditions (ii) and (iii) in Theorem 7.9 implies that D_1 and D_2 are both orthogonal to F. Therefore, as $\mu \in D_1 \cap D_2$, Lemma 5.8 implies that $D_1 = D_2$. Hence, $\lambda_1 = \lambda_2$. □

The following corollary of Theorem 7.9 is equivalent to Theorem 1.2 from the introduction of the paper.

Corollary 7.11 *Let F be a flat geodesic in G and let D be a geodesic in G. Then D has the closest point property with respect to F if and only if D is orthogonal to F.*

References

1. Agler, J., Young, N.J.: The hyperbolic geometry of the symmetrised bidisc. J. Geom. Anal. **14**, 375–403 (2004)
2. Agler, J., Young, N.J.: The complex geodesics of the symmetrised bidisc. Int. J. Math. **17**, 375–391 (2006)
3. Agler, J., Young, N.J.: The magic functions and automorphisms of a domain. Complex Anal. Oper. Theory **2**, 383–404 (2008)
4. Agler, J., Lykova, Z.A., Young, N.J.: Algebraic and geometric aspects of rational Γ-inner functions. Adv. Math. **328**, 133–159 (2018)
5. Agler, J., Lykova, Z.A., Young, N.J.: A geometric characterization of the symmetrized bidisc. J. Math. Anal. Appl. **473**(2), 1377–1413 (2019)
6. Agler, J., Lykova, Z.A., Young, N.J.: Geodesics, retracts, and the extension property in the symmetrized bidisc. Mem. Am. Math. Soc. **258**, 1242 (2019)
7. Ball, J.A., Sau, H.: Rational dilation of tetrablock contractions revisited. J. Funct. Anal. **278**(1), 108275 (2020)
8. Bhattacharyya, T., Pal, S., Roy, S.S.: Dilations of Γ-contractions by solving operator equations. Adv. Math. **230**, 577–606 (2012)
9. Bhattacharyya, T., Das, B.K., Sau, H.: Toeplitz operators on the symmetrized bidisc. Int. Math. Res. Not. (2020). https://doi.org/10.1093/imrn/rnz333
10. Burago, D., Burago, Yu., Ivanov, S.: A Course in Metric Geometry. Graduate Studies in Mathematics, vol. 33. American Mathematical Society, Providence (2001)
11. Jarnicki, M., Pflug, P.: On automorphisms of the symmetrised bidisc. Arch. Math. (Basel) **83**, 264–266 (2004)

12. Jarnicki, M., Pflug, P.: Invariant Distances and Metrics in Complex Analysis, 2nd Extended edn. De Gruyter, Berlin (2013)
13. Kobayashi, S.: Hyperbolic Complex Spaces, Grundlehren der mathematischen Wissenschaften, vol. 318. Springer, Cham (1998)
14. Kosiński, L., McCarthy, J.E.: Norm preserving extensions of bounded holomorphic functions. Trans. Am. Math. Soc. **371**(10), 7243–7257 (2019)
15. Kosiński, L., Zwonek, W.: Nevanlinna–Pick problem and uniqueness of left inverses in convex domains, symmetrized bidisc and tetrablock. J. Geom. Anal. **26**, 1863–1890 (2016)
16. Kosiński, L., Zwonek, W.: Extremal holomorphic maps in special classes of domains, Annali della Scuola Normale Superiore di Pisa. Classe di Scienze **16**(1), 159–182 (2016)
17. Krantz, S.G.: Function Theory of Several Complex Variables. Wiley, New York (1982)
18. Lempert, L.: La métrique de Kobayashi et la représentation des domaines sur la boule. Bull. Soc. Math. Fr. **109**, 427–484 (1981)
19. Pflug, P., Zwonek, W.: Description of all complex geodesics in the symmetrized bidisc. Bull. Lond. Math. Soc. **37**, 575–584 (2005)

Affiliations

Jim Agler[1] · Zinaida Lykova[2] · N. J. Young[2,3]

✉ N. J. Young
  Nicholas.Young@ncl.ac.uk

  Jim Agler
  jagler@ucsd.edu

  Zinaida Lykova
  Zinaida.Lykova@ncl.ac.uk

[1] Department of Mathematics, University of California at San Diego, La Jolla, CA 92103, USA

[2] School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

[3] School of Mathematics, Leeds University, Leeds LS2 9JT, UK