arXiv:1712.00749v2 [math.CV] 5 May 2018

CARATHÉODORY EXTREMAL FUNCTIONS ON THE SYMMETRIZED BIDISC

JIM AGLER, ZINAIDA A. LYKOVA, AND N. J. YOUNG

To Rien Kaashoek in esteem and friendship

Abstract. We show how realization theory can be used to find the solutions of the Carathéodory extremal problem on the symmetrized bidisc

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

We show that, generically, solutions are unique up to composition with automorphisms of the disc. We also obtain formulae for large classes of extremal functions for the Carathéodory problems for tangents of non-generic types.

Introduction

A constant thread in the research of Marinus Kaashoek over several decades has been the power of realization theory applied to a wide variety of problems in analysis. Among his many contributions in this area we mention his monograph [6], written with his longstanding collaborators Israel Gohberg and Harm Bart, which was an early and influential work in the area, and his more recent papers [13, 10]. Realization theory uses explicit formulae for functions in terms of operators on Hilbert space to prove function-theoretic results.

In this paper we continue along the Bart-Gohberg-Kaashoek path by using realization theory to prove results in complex geometry. Specifically, we are interested in the geometry of the *symmetrized bidisc*

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

Date: 5th May 2018.

2010 *Mathematics Subject Classification.* 32A07, 53C22, 54C15, 47A57, 32F45, 47A25, 30E05.

Key words and phrases. Carathéodory extremal functions, symmetrized bidisc, model formulae, realization formulae.

Partially supported by National Science Foundation Grants DMS 1361720 and 1665260, the UK Engineering and Physical Sciences Research Council grant EP/N03242X/1, the London Mathematical Society Grant 41730 and Newcastle University.

a domain in $\mathbb{C}^2$ that has been much studied in the last two decades: see [8, 9, 11, 7, 17, 18, 2], along with many other papers. We shall use realization theory to prove detailed results about the *Carathéodory extremal problem* on G, defined as follows (see [14, 12]).

Consider a domain (that is, a connected open set) Ω in $\mathbb{C}^n$. For domains Ω_1, Ω_2, we denote by $\Omega_2(\Omega_1)$ the set of holomorphic maps from Ω_1 to Ω_2. A point in the complex tangent bundle $T\Omega$ of Ω will be called a *tangent* (to Ω). Thus if $\delta \stackrel{\text{def}}{=} (\lambda, v)$ is a tangent to Ω then $\lambda \in \Omega$ and v is a point in the complex tangent space $T_\lambda\Omega \sim \mathbb{C}^n$ of Ω at λ. We say that δ is a *nondegenerate* tangent if $v \neq 0$. We write $|\cdot|$ for the Poincaré metric on $T\mathbb{D}$:

$$|(z,v)| \stackrel{\text{def}}{=} \frac{|v|}{1-|z|^2} \qquad \text{for } z \in \mathbb{D},\ v \in \mathbb{C}.$$

The *Carathéodory* or *Carathéodory-Reiffen pseudometric* [12] on Ω is the Finsler pseudometric $|\cdot|_{\text{car}}$ on $T\Omega$ defined for $\delta = (\lambda, v) \in T\Omega$ by

$$\begin{aligned} |\delta|_{\text{car}} &\stackrel{\text{def}}{=} \sup_{F \in \mathbb{D}(\Omega)} |F_*(\delta)| \\ &= \sup_{F \in \mathbb{D}(\Omega)} \frac{|D_v F(\lambda)|}{1-|F(\lambda)|^2}. \end{aligned} \tag{0.1}$$

Here F_* is the standard notation for the pushforward of δ by the map F to an element of $T\mathbb{D}$, given by

$$\langle g, F_*(\delta)\rangle = \langle g \circ F, \delta\rangle$$

for any analytic function g in a neighbourhood of $F(\lambda)$.

The *Carathéodory extremal problem* $\operatorname{Car}\delta$ on Ω is to calculate $|\delta|_{\text{car}}$ for a given $\delta \in T\Omega$, and to find the corresponding extremal functions, which is to say, the functions $F \in \mathbb{D}(\Omega)$ for which the supremum in equation (0.1) is attained. We shall also say that F *solves* $\operatorname{Car}\delta$ to mean that F is an extremal function for $\operatorname{Car}\delta$.

For a general domain Ω one cannot expect to find either $|\cdot|_{\text{car}}$ or the corresponding extremal functions explicitly. In a few cases, however, there are more or less explicit formulae for $|\delta|_{\text{car}}$. In particular, when $\Omega = G$, $|\cdot|_{\text{car}}$ is a metric on TG (it is positive for nondegenerate tangents) and the following result obtains [4, Theorem 1.1 and Corollary 4.3]. We use the co-ordinates (s^1, s^2) for a point of G.

Theorem 0.1. *Let δ be a nondegenerate tangent vector in TG. There exists $\omega \in \mathbb{T}$ such that the function in $\mathbb{D}(G)$ given by*

$$\Phi_\omega(s^1, s^2) \stackrel{\text{def}}{=} \frac{2\omega s^2 - s^1}{2 - \omega s^1} \tag{0.2}$$

is extremal for the Carathéodory problem $\operatorname{Car}\delta$ *in* G.

It follows that $|\delta|_{\text{car}}$ can be obtained as the maximum modulus of a fractional quadratic function over the unit circle [4, Corollary 4.4][1]: if $\delta = ((s^1, s^2), v) \in TG$ then

$$\begin{aligned} |\delta|_{\text{car}} &= \sup_{\omega\in\mathbb{T}} |(\Phi_\omega)_*(\delta)| \\ &= \sup_{\omega\in\mathbb{T}} \left| \frac{v_1(1-\omega^2 s^2) - v_2\omega(2-\omega s^1)}{(s^1 - \overline{s^1}s^2)\omega^2 - 2(1-|s^2|^2)\omega + \bar{s} - \overline{s^2}s^1} \right|. \end{aligned}$$

Hence $|\delta|_{\text{car}}$ can easily be calculated numerically to any desired accuracy. In the latter equation we use superscripts (in s^1, s^2) and squares (of ω, $|s^2|$).

The question arises: what are the extremal functions for the problem $\operatorname{Car}\delta$? By Theorem 0.1, there is an extremal function for $\operatorname{Car}\delta$ of the form Φ_ω for some ω in $\mathbb{T}$, but are there others? It is clear that if F is an extremal function for $\operatorname{Car}\delta$ then so is $m \circ F$ for any automorphism m of $\mathbb{D}$, by the invariance of the Poincaré metric on $\mathbb{D}$. We shall say that the solution of $\operatorname{Car}\delta$ is *essentially unique* if, for every pair of extremal functions F_1, F_2 for $\operatorname{Car}\delta$, there exists an automorphism m of $\mathbb{D}$ such that $F_2 = m \circ F_1$.

We show in Theorem 2.1 that, for any nondegenerate tangent $\delta \in TG$, if there is a *unique* ω in $\mathbb{T}$ such that Φ_ω solves $\operatorname{Car}\delta$, then the solution of $\operatorname{Car}\delta$ *is* essentially unique. Indeed, for any point $\lambda \in G$, the solution of $\operatorname{Car}(\lambda, v)$ is essentially unique for generic directions v (Corollary 2.7). We also derive (in Section 3) a parametrization of all solutions of $\operatorname{Car}\delta$ in the special case that δ is tangent to the 'royal variety' $(s^1)^2 = 4s^2$ in G, and in Sections 4 and 5 we obtain large classes of Carathéodory extremals for two other classes of tangents, called *flat* and *purely balanced* tangents.

The question of the essential uniqueness of solutions of $\operatorname{Car}\delta$ in domains including G was studied by L. Kosiński and W. Zwonek in [15]. Their terminology and methods differ from ours; we explain the relation of their Theorem 5.3 to our Theorem 2.1 in Section 6. Incidentally, the authors comment that very little is known about the set of all Carathéodory extremals for a given tangent in a domain. As far as the domain G goes, in this paper we derive a substantial amount of information, even though we do not achieve a complete description of all Carathéodory extremals on G.

---

[1]Unfortunately there is an ω missing in equation (4.7) of [4]. The derivation given there shows that the correct formula is the present one.

The main tool we use is a model formula for analytic functions from G to the closed unit disc $\mathbb{D}^-$ proved in [5] and stated below as Definition 2.2 and Theorem 2.3. Model formulae and realization formulae for a class of functions are essentially equivalent: one can pass back and forth between them by standard methods (algebraic manipulation in one direction, lurking isometry arguments in the other).

1. Five types of tangent

There are certainly nondegenerate tangents $\delta \in TG$ for which the solution of $\operatorname{Car}\delta$ is not essentially unique. Consider, for example, δ of the form

$$\delta = \big((2z, z^2), 2c(1, z)\big)$$

for some $z \in \mathbb{D}$ and nonzero complex c. We call such a tangent *royal*: it is tangent to the 'royal variety'

$$\mathcal{R} \stackrel{\text{def}}{=} \{(2z, z^2) : z \in \mathbb{D}\}$$

in G. By a simple calculation, for any $\omega \in \mathbb{T}$,

$$\Phi_\omega(2z, z^2) = -z, \qquad D_v\Phi_\omega(2z, z^2) = -c,$$

where $v = 2c(1, z)$, so that $\Phi_\omega(2z, z^2)$ and $D_v\Phi_\omega(2z, z^2)$ are independent of ω. It follows from Theorem 0.1 that Φ_ω solves $\operatorname{Car}\delta$ for *all* $\omega \in \mathbb{T}$ and that

$$|\delta|_{\text{car}} = \frac{|D_v\Phi_\omega(2z, z^2)|}{1 - |\Phi_\omega(2z, z^2)|^2} = \frac{|c|}{1 - |z|^2}. \tag{1.1}$$

Now if $\omega_1,\ \omega_2$ are distinct points of $\mathbb{T}$, there is no automorphism m of $\mathbb{D}$ such that $\Phi_{\omega_1} = m \circ \Phi_{\omega_2}$; this is a consequence of the fact that $(2\bar\omega, \bar\omega^2)$ is the unique singularity of Φ_ω in the closure Γ of G. Hence the solution of $\operatorname{Car}\delta$ is not essentially unique.

Similar conclusions hold for another interesting class of tangents, which we call *flat*. These are the tangents of the form

$$(\lambda, v) = \big((\beta + \bar\beta z, z), c(\bar\beta, 1)\big)$$

for some $\beta \in \mathbb{D}$ and $c \in \mathbb{C} \setminus \{0\}$. It is an entertaining calculation to show that

$$|(\lambda, v)|_{\text{car}} = \frac{|D_v\Phi_\omega(\lambda)|}{1 - |\Phi_\omega(\lambda)|^2} = \frac{|c|}{1 - |z|^2} \tag{1.2}$$

for all $\omega \in \mathbb{T}$. Again, the solution to $\operatorname{Car}(\lambda, v)$ is far from being essentially unique.

There are also tangents $\delta \in TG$ such that Φ_ω solves $\operatorname{Car}\delta$ for exactly two values of ω in $\mathbb{T}$; we call these *purely balanced* tangents. They can be described concretely as follows. For any hyperbolic automorphism

m of $\mathbb{D}$ (that is, one that has two fixed points ω_1 and ω_2 in $\mathbb{T}$) let h_m in $G(\mathbb{D})$ be given by

$$h_m(z) = (z + m(z), zm(z))$$

for $z \in \mathbb{D}$. A purely balanced tangent has the form

$$\delta = (h_m(z), ch'_m(z)) \tag{1.3}$$

for some hyperbolic automorphism m of $\mathbb{D}$, some $z \in \mathbb{D}$ and some $c \in \mathbb{C} \setminus \{0\}$. It is easy to see that, for $\omega \in \mathbb{T}$, the composition $\Phi_\omega \circ h_m$ is a rational inner function of degree at most 2 and that the degree reduces to 1 precisely when ω is either $\bar{\omega}_1$ or $\bar{\omega}_2$. Thus, for these two values of ω (and only these), $\Phi_\omega \circ h_m$ is an automorphism of $\mathbb{D}$. It follows that Φ_ω solves $\operatorname{Car}\delta$ if and only if $\omega = \bar{\omega}_1$ or $\bar{\omega}_2$.

A fourth type of tangent, which we call *exceptional*, is similar to the purely balanced type, but differs in that the hyperbolic automorphism m of $\mathbb{D}$ is replaced by a *parabolic* automorphism, that is, an automorphism m of $\mathbb{D}$ which has a single fixed point ω_1 in $\mathbb{T}$, which has multiplicity 2. The same argument as in the previous paragraph shows that Φ_ω solves the Carathéodory problem if and only if $\omega = \bar{\omega}_1$.

The fifth and final type of tangent is called *purely unbalanced*. It consists of the tangents $\delta = (\lambda, v) \in TG$ such that Φ_ω solves $\operatorname{Car}\delta$ for a unique value e^{it_0} of ω in $\mathbb{T}$ and

$$\frac{d^2}{dt^2} \frac{|D_v \Phi_{\mathrm{e}^{it}}(\lambda)|}{1 - |\Phi_{\mathrm{e}^{it}}(\lambda)|^2}\bigg|_{t=t_0} < 0. \tag{1.4}$$

The last inequality distinguishes purely unbalanced from exceptional tangents – the left hand side of equation (1.4) is equal to zero for exceptional tangents.

The five types of tangent are discussed at length in our paper [2]. We proved [2, Theorem 3.6] a 'pentachotomy theorem', which states that every nondegenerate tangent in TG is of exactly one of the above five types. We also give, for a representative tangent of each type, a cartoon showing the unique complex geodesic in G touched by the tangent [2, Appendix B].

It follows trivially from Theorem 0.1 that, for every nondegenerate tangent $\delta \in TG$, either

(1) there exists a unique $\omega \in \mathbb{T}$ such that Φ_ω solves $\operatorname{Car}\delta$, or
(2) there exist at least two values of ω in $\mathbb{T}$ such that Φ_ω solves $\operatorname{Car}\delta$.

The above discussion shows that Case (1) obtains for purely unbalanced and exceptional tangents, while Case (2) holds for royal, flat and purely balanced tangents. For the purpose of this paper, the message to be drawn is that Case (1) is generic in the following sense. Consider a point $\lambda \in G$. Each tangent v in $T_\lambda G$ has a 'complex direction' $\mathbb{C}v$, which is a one-dimensional subspace of $\mathbb{C}^2$, or in other words, a point of the projective space CP^2. The directions corresponding to the royal (if any) and flat tangents at λ are just single points in CP^2, while, from the constructive nature of the expression (1.3) for a purely balanced tangent, it is easy to show that there is a smooth one-real-parameter curve of purely balanced directions (see [1, Section 1]). It follows that the set of directions $\mathbb{C}v \in \mathrm{CP}^2$ for which a unique Φ_ω solves $\mathrm{Car}\,\delta$ contains a dense open set in CP^2. To summarise:

Proposition 1.1. *For every $\lambda \in G$ there exists a dense open set V_λ in CP^2 such that whenever $\mathbb{C}v \in V_\lambda$, there exists a unique $\omega \in \mathbb{T}$ such that Φ_ω solves $\mathrm{Car}(\lambda, v)$.*

2. Tangents with a unique extremal Φ_ω

In Section 1 we discussed extremal functions of the special form Φ_ω, $\omega \in \mathbb{T}$, for the Carathéodory problem in G. However, there is no reason to expect that the Φ_ω will be the only extremal functions. For example, if $\delta = (\lambda, v)$ is a nondegenerate tangent and $\Phi_{\omega_1}, \dots, \Phi_{\omega_k}$ all solve $\mathrm{Car}\,\delta$, then one can generate a large class of other extremal functions as follows. Choose an automorphism m_j of $\mathbb{D}$ such that $m_j \circ \Phi_{\omega_j}(\lambda) = 0$ and $D_v(m_j \circ \Phi_{\omega_j})(\lambda_j) > 0$ for $j = 1, \dots, k$. Then each $m_j \circ \Phi_{\omega_j}$ solves $\mathrm{Car}\,\delta$, and so does any convex combination of them.

Nevertheless, if there is a *unique* $\omega \in \mathbb{T}$ such that Φ_ω is extremal for $\mathrm{Car}\,\delta$ then the solution of $\mathrm{Car}\,\delta$ *is* essentially unique.

Theorem 2.1. *Let δ be a nondegenerate tangent in G such that Φ_ω solves $\mathrm{Car}\,\delta$ for a unique value of ω in $\mathbb{T}$. If ψ solves $\mathrm{Car}\,\delta$ then there exists an automorphism m of $\mathbb{D}$ such that $\psi = m \circ \Phi_\omega$.*

For the proof recall the following model formula [5, Definition 2.1 and Theorem 2.2].

Definition 2.2. *A G-*model *for a function φ on G is a triple $(\mathcal{M}, T, u)$ where $\mathcal{M}$ is a separable Hilbert space, T is a contraction acting on $\mathcal{M}$ and $u : G \to \mathcal{M}$ is an analytic map such that, for all $s, t \in G$,*

$$1 - \overline{\varphi(t)}\varphi(s) = \langle (1 - t_T^* s_T) u(s), u(t) \rangle_{\mathcal{M}} \tag{2.1}$$

where, for $s \in G$,

$$s_T \stackrel{\text{def}}{=} (2s^2 T - s^1)(2 - s^1 T)^{-1}.$$

A G-model $(\mathcal{M}, T, u)$ is unitary *if T is a unitary operator on $\mathcal{M}$.*

For any domain Ω we define the *Schur class* $\mathcal{S}(\Omega)$ to be the set of holomorphic maps from Ω to the closed unit disc $\mathbb{D}^-$.

Theorem 2.3. *Let φ be a function on G. The following three statements are equivalent.*

(1) $\varphi \in \mathcal{S}(G)$*;*
(2) φ *has a G-model;*
(3) φ *has a unitary G-model* $(\mathcal{M}, T, u)$*.*

From a G-model of a function $\varphi \in \mathcal{S}(G)$ one may easily proceed by means of a standard lurking isometry argument to a realization formula

$$\varphi(s) = A + Bs_T(1 - Ds_T)^{-1}C, \quad \text{all } s \in G,$$

for φ, where $ABCD$ is a contractive or unitary colligation on $\mathbb{C} \oplus \mathcal{M}$. However, for the present purpose it is convenient to work directly from the G-model.

We also require a long-established fact about G [4], related to the fact that the Carathéodory and Kobayashi metrics on TG coincide.

Lemma 2.4. *If δ is a nondegenerate tangent to G and φ solves* $\operatorname{Car}\delta$ *then there exists k in $G(\mathbb{D})$ such that $\varphi \circ k = \mathrm{id}_{\mathbb{D}}$. Moreover, if ψ is any solution of* $\operatorname{Car}\delta$ *then $\psi \circ k$ is an automorphism of $\mathbb{D}$.*

We shall need some minor measure-theoretic technicalities.

Lemma 2.5. *Let Y be a set and let*

$$A : \mathbb{T} \times Y \times Y \to \mathbb{C}$$

be a map such that

(1) *$A(\cdot, z, w)$ is continuous on $\mathbb{T}$ for every $z, w \in Y$;*
(2) *$A(\eta, \cdot, \cdot)$ is a positive kernel on Y for every $\eta \in \mathbb{T}$.*

Let $\mathcal{M}$ be a separable Hilbert space, let T be a unitary operator on $\mathcal{M}$ with spectral resolution

$$T = \int_{\mathbb{T}} \eta \, \mathrm{d}E(\eta)$$

and let $v : Y \to \mathcal{M}$ be a mapping. Let

$$C(z, w) = \int_{\mathbb{T}} A(\eta, z, w) \; \langle \mathrm{d}E(\eta)v(z), v(w)\rangle \tag{2.2}$$

for all $z, w \in Y$. Then C is a positive kernel on Y.

Proof. Consider any finite subset $\{z_1, \dots, z_N\}$ of Y. We must show that the $N \times N$ matrix

$$\left[C(z_i, z_j)\right]_{i,j=1}^{N}$$

is positive.

Since $A(\cdot, z_i, z_j)$ is continuous on $\mathbb{T}$ for each i and j, we may approximate the $N \times N$-matrix-valued function $[A(\cdot, z_i, z_j)]$ uniformly on $\mathbb{T}$ by integrable simple functions of the form

$$[f_{ij}] = \sum_\ell b^\ell \chi_{\tau_\ell}$$

for some $N \times N$ matrices b^ℓ and Borel sets τ_ℓ, where χ denotes 'characteristic function'. Moreover we may do this in such a way that each b^ℓ is a value $[A(\eta, z_i, z_j)]$ for some $\eta \in \mathbb{T}$, hence is positive. Then

$$\left[\int_\tau f_{ij}(\eta)\ \langle \mathrm{d}E(\eta) z_i, z_j\rangle\right]_{i,j=1}^N = \sum_\ell b^\ell * \left[\langle E(\tau_\ell) v_i, v_j\rangle\right]_{i,j=1}^N \tag{2.3}$$

where $*$ denotes the Schur (or Hadamard) product of matrices. Since the matrix $\left[\langle E(\tau_\ell) v_i, v_j\rangle\right]$ is positive and the Schur product of positive matrices is positive, every approximating sum of the form (2.3) is positive, and hence the integral in equation (2.2) is a positive matrix. $\square$

Lemma 2.6. *For $i, j = 1, 2$ let $a_{ij} : \mathbb{T} \to \mathbb{C}$ be continuous and let each a_{ij} have only finitely many zeros in $\mathbb{T}$. Let ν_{ij} be a complex-valued Borel measure on $\mathbb{T}$ such that, for every Borel set τ in $\mathbb{T}$,*

$$\left[\nu_{ij}(\tau)\right]_{i,j=1}^2 \geq 0.$$

Let X be a Borel subset of $\mathbb{T}$ and suppose that

$$\left[a_{ij}(\eta)\right]_{i,j=1}^2 \quad \text{is positive and of rank 2 for all } \eta \in X.$$

Let

$$C = \left[c_{ij}\right]_{i,j=1}^2$$

where

$$c_{ij} = \int_X a_{ij}(\eta)\, \mathrm{d}\nu_{ij}(\eta) \qquad \text{for } i, j = 1, 2.$$

If $\operatorname{rank} C \leq 1$ then either $c_{11} = 0$ or $c_{22} = 0$.

Proof. By hypothesis the set

$$Z \stackrel{\text{def}}{=} \bigcup_{i,j=1}^2 \{\eta \in \mathbb{T} : a_{ij}(\eta) = 0\}$$

is finite.

Exactly as in the proof of Lemma 2.5, for any Borel set τ in $\mathbb{T}$,

$$\left[\int_\tau a_{ij}\, \mathrm{d}\nu_{ij}\right]_{ij=1}^2 \geq 0. \tag{2.4}$$

Suppose that C has rank at most 1 but c_{11} and c_{22} are both nonzero. Then there exists a nonzero 2×1 matrix $c = [c_1\, c_2]^T$ such that $C = cc^*$ for $i, j = 1, 2$ and $c_1,\ c_2$ are nonzero.

For any Borel set $\tau \subset X$,

$$\left[\int_\tau a_{ij}\,\mathrm{d}\nu_{ij}\right] \le \left[\int_\tau + \int_{X\setminus\tau} a_{ij}\,\mathrm{d}\nu_{ij}\right] = \left[\int_X a_{ij}\,\mathrm{d}\nu_{ij}\right] = C = cc^*.$$

Consequently there exists a unique $\mu(\tau) \in [0, 1]$ such that

$$\left[\int_\tau a_{ij}\,\mathrm{d}\nu_{ij}\right] = \mu(\tau)C. \tag{2.5}$$

It is easily seen that μ is a Borel probability measure on X. Note that if $\eta \in Z$, say $a_{ij}(\eta) = 0$, then on taking $\tau = \{\eta\}$ in equation (2.5), we deduce that

$$\mu(\{\eta\})c_i\bar{c}_j = 0.$$

Since c_1, c_2 are nonzero, it follows that $\mu(\{\eta\}) = 0$. Hence $\mu(Z) = 0$.

Equation (2.5) states that μ is absolutely continuous with respect to ν_{ij} on X and the Radon-Nikodym derivative is given by

$$c_i\bar{c}_j\frac{\mathrm{d}\mu}{\mathrm{d}\nu_{ij}} = a_{ij}$$

for $i, j = 1, 2$. Hence, on $X \setminus Z$,

$$\mathrm{d}\nu_{ij} = \frac{c_i\bar{c}_j}{a_{ij}}\mathrm{d}\mu, \qquad i, j = 1, 2. \tag{2.6}$$

Pick a compact subset K of $X \setminus Z$ such that $\mu(K) > 0$. This is possible, since $\mu(X\setminus Z) = 1$ and Borel measures on $\mathbb{T}$ are automatically regular. By compactness, there exists a point $\eta_0 \in K$ such that, for every open neighbourhood U of η_0,

$$\mu(U \cap K) > 0.$$

Notice that, for $\eta \in \mathbb{T} \setminus Z$,

$$\det\left[\frac{c_i\bar{c}_j}{a_{ij}(\eta)}\right]_{i,j=1}^2 = -\frac{|c_1c_2|^2\det\left[a_{ij}(\eta)\right]}{a_{11}(\eta)a_{22}(\eta)|a_{12}(\eta)|^2} < 0.$$

Thus $[c_i\bar{c}_j a_{ij}(\eta_0)^{-1}]$ has a negative eigenvalue. Therefore there exists a unit vector $x \in \mathbb{C}^2$, an $\varepsilon > 0$ and an open neighourhood U of η_0 in $\mathbb{T}$ such that

$$\left\langle \left[c_i\bar{c}_j a_{ij}(\eta)^{-1}\right] x, x\right\rangle < -\varepsilon$$

for all $\eta \in U$. We then have

$$\begin{aligned} \left\langle \left[\nu_{ij}(U\cap K)\right] x, x \right\rangle &= \left\langle \int_{U\cap K} \left[c_i \bar{c}_j a_{ij}(\eta)^{-1}\right] \mathrm{d}\mu(\eta) x, x \right\rangle \\ &= \int_{U\cap K} \left\langle \left[c_i \bar{c}_j a_{ij}(\eta)^{-1}\right] x, x \right\rangle \mathrm{d}\mu(\eta) \\ &< -\varepsilon \mu(U\cap K) \\ &< 0. \end{aligned}$$

This contradicts the positivity of the matricial measure $\left[\nu_{ij}\right]$. Hence either $c_1 = 0$ or $c_2 = 0$.

□

Proof of Theorem 2.1. Let δ be a nondegenerate tangent to G such that Φ_ω is the unique function from the collection $\{\Phi_\eta\}_{\eta\in\mathbb{T}}$ that solves $\operatorname{Car}\delta$. Let ψ be a solution of $\operatorname{Car}\delta$. We must find an automorphism m of $\mathbb{D}$ such that $\psi = m \circ \Phi_\omega$.

By Lemma 2.4, there exists k in $G(\mathbb{D})$ such that

$$\Phi_\omega \circ k = \mathrm{id}_{\mathbb{D}}, \tag{2.7}$$

and moreover, the function

$$m \stackrel{\text{def}}{=} \psi \circ k \tag{2.8}$$

is an automorphism of $\mathbb{D}$. Let

$$\varphi = m^{-1} \circ \psi. \tag{2.9}$$

Then

$$\varphi \circ k = m^{-1} \circ \psi \circ k = m^{-1} \circ m = \mathrm{id}_{\mathbb{D}}. \tag{2.10}$$

By Theorem 2.3, there is a unitary G-model $(\mathcal{M}, T, u)$ for φ. By the Spectral Theorem for unitary operators, there is a spectral measure $E(.)$ on $\mathbb{T}$ with values in $\mathcal{B}(\mathcal{M})$ such that

$$T = \int_{\mathbb{T}} \eta \, \mathrm{d}E(\eta).$$

Thus, for $s \in G$,

$$s_T = (2s^2 T - s^1)(2 - s^1 T)^{-1} = \int_{\mathbb{T}} \Phi_\eta(s) \, \mathrm{d}E(\eta).$$

Therefore, for all $s, t \in G$,

$$\begin{aligned} 1 - \overline{\varphi(t)}\varphi(s) &= \langle (1 - t_T^* s_T) u(s), u(t) \rangle_{\mathcal{M}} \\ &= \int_{\mathbb{T}} \left(1 - \overline{\Phi_\eta(t)}\Phi_\eta(s)\right) \langle \mathrm{d}E(\eta) u(s), u(t) \rangle_{\mathcal{M}} . \end{aligned} \tag{2.11}$$

Consider $z, w \in \mathbb{D}$, put $s = k(z)$, $t = k(w)$ in equation (2.11). Invoke equation (2.10) and divide equation (2.11) through by $1-\bar{w}z$ to obtain, for $z, w \in \mathbb{D}$,

$$\begin{aligned} 1 &= \int_{\{\omega\}} + \int_{\mathbb{T}\setminus\{\omega\}} \frac{1-\overline{\Phi_\eta \circ k(w)}\Phi_\eta \circ k(z)}{1-\bar{w}z} \langle \mathrm{d}E(\eta) u\circ k(z), u \circ k(w)\rangle \\ &= I_1 + I_2 \end{aligned} \tag{2.12}$$

where

$$\begin{aligned} I_1(z,w) &= \langle E(\{\omega\}) u\circ k(z), u\circ k(w)\rangle, \\ I_2(z,w) &= \int_{\mathbb{T}\setminus\{\omega\}} \frac{1-\overline{\Phi_\eta \circ k(w)}\Phi_\eta \circ k(z)}{1-\bar{w}z} \langle \mathrm{d}E(\eta) u\circ k(z), u \circ k(w)\rangle . \end{aligned} \tag{2.13}$$

The left hand side 1 of equation (2.12) is a positive kernel of rank one on $\mathbb{D}$, and I_1 is also a positive kernel. The integrand in I_2 is a positive kernel on $\mathbb{D}$ for each $\eta \in \mathbb{T}$, by Pick's theorem, since $\Phi_\eta \circ k$ is in the Schur class. Hence, by Lemma 2.5, I_2 is also a positive kernel on $\mathbb{D}$. Since $I_1 + I_2$ has rank 1, it follows that I_2 has rank at most 1 as a kernel on $\mathbb{D}$.

By hypothesis, Φ_η does *not* solve Car δ for any $\eta \in \mathbb{T}\setminus\{\omega\}$. Therefore $\Phi_\eta \circ k$ is a Blaschke product of degree 2, and consequently, for any choice of distinct points z_1, z_2 in $\mathbb{D}$, the 2×2 matrix

$$\left[a_{ij}(\eta)\right]_{i,j=1}^2 \stackrel{\text{def}}{=} \left[\frac{1-\overline{\Phi_\eta \circ k(z_i)}\Phi_\eta \circ k(z_j)}{1-\bar{z}_i z_j}\right]_{i,j=1}^2 \tag{2.14}$$

is a positive matrix of rank 2 for every $\eta \in \mathbb{T}\setminus\{\omega\}$. In particular, $a_{11}(\eta) > 0$ for all $\eta \in \mathbb{T}\setminus\{\omega\}$.

Moreover, each a_{ij} has only finitely many zeros in $\mathbb{T}$, as may be seen from the fact that a_{ij} is a ratio of trigonometric polynomials in η. To be explicit, if we temporarily write $k = (k^1, k^2) : \mathbb{D} \to G$, then equation (2.14) expands to $a_{ij}(\eta) = P(\eta)/Q(\eta)$ where

$$\begin{aligned} P(\eta) &= 4\left(1-\overline{k^2(z_i)}k^2(z_j)\right) - 2\eta\left(k^1(z_j) - \overline{k^1(z_i)}k^2(z_j)\right) \\ &\qquad - 2\bar{\eta}\left(\overline{k^1(z_i)} - \overline{k^2(z_i)}k^1(z_j)\right), \\ Q(\eta) &= (1-\bar{z}_i z_j)(2-\eta k^1(z_i))^-(2-\eta k^1(z_j)). \end{aligned}$$

Let

$$\nu_{ij} = \langle E(\cdot) u \circ k(z_i), u\circ k(z_j)\rangle .$$

Clearly $[\nu_{ij}(\tau)] \geq 0$ for every Borel subset τ of $\mathbb{T} \setminus \{\omega\}$. By definition (2.13),

$$I_2(z_i, z_j) = \int_{\mathbb{T}\setminus\{\omega\}} a_{ij} \, \mathrm{d}\nu_{ij}$$

for $i, j = 1, 2$. Moreover, by equation (2.12),

$$[I_2(z_i, z_j)] \leq [I_1(z_i, z_j)] + [I_2(z_i, z_j)] = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

It follows that

$$\left[\int_{\mathbb{T}\setminus\{\omega\}} a_{ij} \, \mathrm{d}\nu_{ij}\right] = [I_2(z_i, z_j] = \kappa \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{2.15}$$

for some $\kappa \in [0, 1]$. We may now apply Lemma 2.6 with $X = \mathbb{T} \setminus \{\omega\}$ to deduce that $\kappa = 0$ and hence $I_2(z_i, z_j) = 0$. In particular,

$$0 = I_2(z_1, z_1) = \int_{\mathbb{T}\setminus\{\omega\}} a_{11} \, \mathrm{d}\nu_{11}.$$

Since $a_{11} > 0$ on $\mathbb{T} \setminus \{\omega\}$, it follows that $\nu_{11}(\mathbb{T} \setminus \{\omega\}) = 0$, which is to say that

$$E(\mathbb{T} \setminus \{\omega\}) u \circ k(z_1) = 0. \tag{2.16}$$

Since z_1, z_2 were chosen arbitrarily in $\mathbb{T} \setminus \{\omega\}$, we have $I_2 \equiv 0$ and therefore, by equation (2.12),

$$1 = I_1 = \langle E(\{\omega\}) u \circ k(z), u \circ k(w) \rangle \tag{2.17}$$

for all $z, w \in \mathbb{D}$. It follows that

$$\|E(\{\omega\}) u \circ k(z) - E(\{\omega\}) u \circ k(w)\|^2 = 0$$

for all z, w, and hence that there exists a unit vector $x \in \mathcal{M}$ such that

$$E(\{\omega\}) u \circ k(z) = x$$

for all $z \in \mathbb{D}$.

In equation (2.11), choose $t = k(w)$ for some $w \in \mathbb{D}$. Since $\Phi_\omega \circ k = \mathrm{id}_{\mathbb{D}}$, we have for all $s \in G$,

$$\begin{aligned} 1 - \bar{w}\varphi(s) &= 1 - \overline{\varphi \circ k(w)}\varphi(s) \\ &= \int_{\{\omega\}} + \int_{\mathbb{T}\setminus\{\omega\}} \left(1 - \overline{\Phi_\eta \circ k(w)} \Phi_\eta(s)\right) \langle \mathrm{d}E(\eta) u(s), u \circ k(w) \rangle \\ &= (1 - \bar{w}\Phi_\omega(s)) \langle u(s), x \rangle + \\ &\qquad \int_{\mathbb{T}\setminus\{\omega\}} \left(1 - \overline{\Phi_\eta \circ k(w)} \Phi_\eta(s)\right) \langle \mathrm{d}E(\eta) u(s), u \circ k(w) \rangle . \end{aligned}$$

In view of equation (2.16), the scalar spectral measure in the second term on the right hand side is zero on $\mathbb{T} \setminus \{\omega\}$. Hence the integral is zero, and so, for all $s \in G$ and $w \in \mathbb{D}$,

$$1 - \bar{w}\varphi(s) = (1 - \bar{w}\Phi_\omega(s)) \langle u(s), x \rangle . \tag{2.18}$$

Put $w = 0$ to deduce that

$$\langle u(s), x \rangle = 1$$

for all $s \in G$, then equate coefficients of $\bar{w}$ to obtain $\varphi = \Phi_\omega$. Hence, by equation (2.8),

$$\psi = m \circ \varphi = m \circ \Phi_\omega$$

as required. □

On combining Theorem 2.1 and Proposition 1.1 we obtain the statement in the abstract.

Corollary 2.7. *Let $\lambda \in G$. For a generic direction $\mathbb{C}v$ in CP^2, the solution of the Carathéodory problem $\mathrm{Car}(\lambda, v)$ is essentially unique.*

It will sometimes be useful in the sequel to distinguish a particular Carathéodory extremal function from a class of functions that are equivalent up to composition with automorphisms of $\mathbb{D}$. Consider any tangent $\delta \in TG$ and any solution φ of $\mathrm{Car}\,\delta$. The functions $m \circ \varphi$, with m an automorphism of $\mathbb{D}$, also solve $\mathrm{Car}\,\delta$, and among them there is exactly one that has the property

$$m \circ \varphi(\lambda) = 0 \quad \text{and} \quad D_v(m \circ \varphi)(\lambda) > 0,$$

or equivalently,

$$(m \circ \varphi)_*(\delta) = (0, |\delta|_{\mathrm{car}}). \tag{2.19}$$

We shall say that φ is *well aligned at δ* if $\varphi_*(\delta) = (0, |\delta|_{\mathrm{car}})$.

With this terminology the following is a re-statement of Theorem 2.1.

Corollary 2.8. *If δ is a nondegenerate tangent in G such that Φ_ω solves $\mathrm{Car}\,\delta$ for a unique value of ω in $\mathbb{T}$ then there is a unique well-aligned solution of $\mathrm{Car}\,\delta$. It is expressible as $m \circ \Phi_\omega$ for some automorphism m of $\mathbb{D}$.*

3. Royal tangents

At the opposite extreme from the tangents studied in the last section are the royal tangents to G. Recall that these have the form

$$\delta = \big((2z, z^2), 2c(1, z)\big) \tag{3.1}$$

for some $z \in \mathbb{D}$ and nonzero complex number c. As we observed in Section 1,

$$|\delta|_{\mathrm{car}} = \frac{|c|}{1-|z|^2}$$

and *all* $\Phi_\omega, \omega \in \mathbb{T}$, solve $\mathrm{Car}\,\delta$. In this section we shall describe *all* extremal functions for $\mathrm{Car}\,\delta$ for royal tangents δ, not just those of the form Φ_ω.

Theorem 3.1. *Let $\delta \in TG$ be the royal tangent*

$$\delta = \big((2z, z^2), 2c(1,z)\big) \tag{3.2}$$

for some $z \in \mathbb{D}$ and $c \in \mathbb{C} \setminus \{0\}$. A function $\varphi \in \mathbb{D}(G)$ solves $\mathrm{Car}\,\delta$ *if and only if there exists an automorphism m of $\mathbb{D}$ and $\Psi \in \mathcal{S}(G)$ such that, for all $s \in G$,*

$$\varphi(s) = m\left(\tfrac{1}{2}s^1 + \tfrac{1}{4}((s^1)^2 - 4s^2)\frac{\Psi(s)}{1-\frac{1}{2}s^1\Psi(s)}\right). \tag{3.3}$$

Proof. We shall lift the problem $\mathrm{Car}\,\delta$ to a Carathéodory problem on the bidisc $\mathbb{D}^2$, where we can use the results of [3] on the Nevanlinna-Pick problem on the bidisc.

Let $\pi : \mathbb{D}^2 \to G$ be the 'symmetrization map',

$$\pi(\lambda^1, \lambda^2) = (\lambda^1 + \lambda^2, \lambda^1\lambda^2)$$

and let $k : \mathbb{D} \to \mathbb{D}^2$ be given by $k(\zeta) = (\zeta, \zeta)$ for $\zeta \in \mathbb{D}$.

Consider the royal tangent δ of equation (3.2) and let

$$\delta_{zc} = ((z,z),(c,c)) \in T\mathbb{D}^2.$$

Observe that

$$\pi'(\lambda) = \begin{bmatrix} 1 & 1 \\ \lambda^2 & \lambda^1 \end{bmatrix},$$

and so

$$\pi_*(\delta_{zc}) = (\pi(z,z), \pi'(z,z)(c,c)) = \big((2z, z^2), 2c(1,z)\big) = \delta, \tag{3.4}$$

while

$$k_*((z,c)) = (k(z), k'(z)c) = ((z,z),(c,c)) = \delta_{zc}.$$

Consider any $\varphi \in \mathbb{D}(G)$. Figure 1 illustrates the situation.

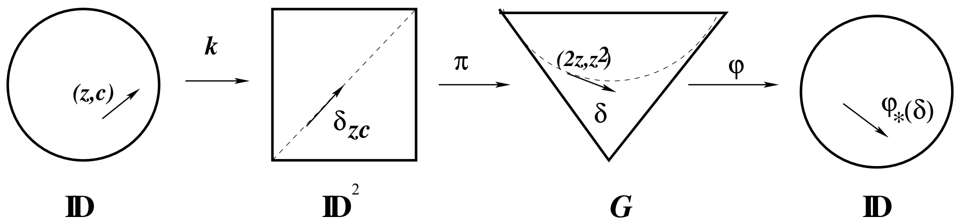

FIGURE 1

It is known that every Carathéodory problem on the bidisc is solved by one of the two co-ordinate functions $F_j(\lambda) = \lambda^j$ for $j = 1$ or 2 (for a proof see, for example, [2, Theorem 2.3]). Thus

$$\begin{aligned} |\delta_{zc}|^{\mathbb{D}^2}_{\mathrm{car}} &= \max_{j=1,2} \frac{|D_{(c,c)}F_j(z,z)|}{1-|F_j(z,z)|^2} \\ &= \frac{|c|}{1-|z|^2} \\ &= |\delta|_{\mathrm{car}}. \end{aligned}$$

Here of course the superscript $\mathbb{D}^2$ indicates the Carathéodory extremal problem on the bidisc.

Hence, for $\varphi \in \mathbb{D}(G)$,

$$\begin{aligned} \varphi\circ\pi \text{ solves } \operatorname{Car}\delta_{zc} &\iff |(\varphi\circ\pi)_*(\delta_{zc})| = \frac{|c|}{1-|z|^2} \\ &\iff |\varphi_*\circ\pi_*(\delta_{zc})| = \frac{|c|}{1-|z|^2} \qquad \text{by the chain rule} \\ &\iff |\varphi_*(\delta)| = \frac{|c|}{1-|z|^2} \qquad \text{by equation (3.4)} \\ &\iff \varphi \text{ solves } \operatorname{Car}\delta. \end{aligned} \tag{3.5}$$

Next observe that a function $\psi \in \mathbb{D}(\mathbb{D}^2)$ solves $\operatorname{Car}\delta_{zc}$ if and only if $\psi\circ k$ is an automorphism of $\mathbb{D}$. For if $\psi\circ k$ is an automorphism of $\mathbb{D}$ then it satisfies

$$|(z,c)| = |(\psi\circ k)_*(z,c)| = |\psi_*\circ k_*(z,c)| = |\psi_*(\delta_{zc})|,$$

which is to say that ψ solves $\operatorname{Car}\delta_{zc}$. Conversely, if ψ solves $\operatorname{Car}\delta_{zc}$ then $\psi\circ k$ is an analytic self-map of $\mathbb{D}$ that preserves the Poincaré metric of a nondegenerate tangent to $\mathbb{D}$, and is therefore (by the Schwarz-Pick lemma) an automorphism of $\mathbb{D}$. On combining this observation with equivalence (3.5) we deduce that

$$\varphi \text{ solves } \operatorname{Car}\delta \iff \begin{array}{l}\text{there exists an automorphism } m \text{ of } \mathbb{D}\\ \text{such that } m^{-1}\circ\varphi\circ\pi\circ k = \mathrm{id}_{\mathbb{D}}.\end{array} \tag{3.6}$$

For a function $f \in \mathbb{D}(\mathbb{D}^2)$, it is easy to see that $f \circ k = \mathrm{id}_{\mathbb{D}}$ if and only if f solves the Nevanlinna-Pick problem

$$(0,0) \mapsto 0, \qquad (\tfrac{1}{2}, \tfrac{1}{2}) \mapsto \tfrac{1}{2}. \tag{3.7}$$

See [3, Subsection 11.5] for the Nevanlinna-Pick problem in the bidisc. Hence

$$\varphi \text{ solves } \operatorname{Car}\delta \iff \begin{array}{l}\text{there exists an automorphism } m \text{ of } \mathbb{D} \text{ such that}\\ m^{-1}\circ\varphi\circ\pi \text{ solves the Nevanlinna-Pick problem (3.7).}\end{array} \tag{3.8}$$

In [3, Subsection 11.6] Agler and McCarthy use realization theory to show the following.

A function $f \in \mathcal{S}(\mathbb{D}^2)$ satisfies the interpolation conditions

$$f(0,0) = 0, \qquad f(\tfrac{1}{2}, \tfrac{1}{2}) = \tfrac{1}{2} \tag{3.9}$$

if and only if there exist $t \in [0,1]$ and Θ in the Schur class of the bidisc such that, for all $\lambda \in \mathbb{D}^2$,

$$f(\lambda) = t\lambda^1 + (1-t)\lambda^2 + t(1-t)(\lambda^1 - \lambda^2)^2 \frac{\Theta(\lambda)}{1 - [(1-t)\lambda^1 + t\lambda^2]\Theta(\lambda)}. \tag{3.10}$$

Inspection of the formula (3.10) reveals that f is symmetric if and only if $t = \frac{1}{2}$ and Θ is symmetric. Hence the symmetric functions in $\mathcal{S}(\mathbb{D}^2)$ that satisfy the conditions (3.9) are those given by

$$f(\lambda) = \tfrac{1}{2}\lambda^1 + \tfrac{1}{2}\lambda^2 + \tfrac{1}{4}(\lambda^1 - \lambda^2)^2 \frac{\Theta(\lambda)}{1 - \frac{1}{2}(\lambda^1 + \lambda^2)\Theta(\lambda)} \tag{3.11}$$

for some symmetric $\Theta \in \mathcal{S}(\mathbb{D}^2)$. Such a Θ induces a unique function $\Psi \in \mathcal{S}(G)$ such that $\Theta = \Psi \circ \pi$, and we may write the symmetric solutions f of the problem (3.9) in the form $f = \tilde{f} \circ \pi$ where, for all $s = (s^1, s^2)$ in G,

$$\tilde{f}(s) = \tfrac{1}{2}s^1 + \tfrac{1}{4}((s^1)^2 - 4s^2)\frac{\Psi(s)}{1 - \frac{1}{2}s^1\Psi(s)}. \tag{3.12}$$

Let φ solve $\operatorname{Car}\delta$. By the equivalence (3.8), there exists an automorphism m of $\mathbb{D}$ such that $m^{-1}\circ\varphi\circ\pi$ solves the Nevanlinna-Pick problem

(3.7). Clearly $m^{-1}\circ\varphi\circ\pi$ is symmetric. Hence there exists $\Psi\in\mathcal{S}(G)$ such that, for all $\lambda\in\mathbb{D}^2$,

$$m^{-1}\circ\varphi(s)=\tfrac{1}{2}s^1+\tfrac{1}{4}((s^1)^2-4s^2)\frac{\Psi(s)}{1-\frac{1}{2}s^1\Psi(s)}. \tag{3.13}$$

Thus φ is indeed given by the formula (3.3).

Conversely, suppose that for some automorphism m of $\mathbb{D}$ and $\Psi\in\mathcal{S}(G)$, a function φ is defined by equation (3.3). Let $f=m^{-1}\circ\varphi\circ\pi$ Then f is given by the formula (3.11), where $\Theta=\Psi\circ\pi$. Hence f is a symmetric function that satisfies the interpolation conditions (3.9). By the equivalence (3.8), φ solves $\operatorname{Car}\delta$. □

4. Flat tangents

In this section we shall give a description of a large class of Carathéodory extremals for a flat tangent. Recall that a flat tangent has the form

$$\delta=\big((\beta+\bar\beta z,z),c(\bar\beta,1)\big) \tag{4.1}$$

for some $z\in\mathbb{D}$ and $c\neq 0$, where $\beta\in\mathbb{D}$. Such a tangent touches the 'flat geodesic'

$$F_\beta\overset{\text{def}}{=}\{(\beta+\bar\beta w,w):w\in\mathbb{D}\}.$$

The description depends on a remarkable property of sets of the form $\mathcal{R}\cup F_\beta$, $\beta\in\mathbb{D}$: they have the norm-preserving extension property in G [2, Theorem 10.1]. That is, if g is any bounded analytic function on the variety $\mathcal{R}\cup F_\beta$, then there exists an analytic function $\tilde g$ on G such that $g=\tilde g|\mathcal{R}\cup F_\beta$ and the supremum norms of g and $\tilde g$ coincide. Indeed, the proof of [2, Theorem 10.1] gives an explicit formula for one such $\tilde g$ in terms of a Herglotz-type integral. Let us call the norm-preserving extension $\tilde g$ of g constructed in [2, Chapter 10] the *special extension* of g to G.

It is a simple calculation to show that $\mathcal{R}$ and F_β have a single point in common.

By equation (1.2), for δ in equation (4.1)

$$|\delta|_{\text{car}}=\frac{|c|}{1-|z|^2}.$$

Theorem 4.1. *Let δ be the flat tangent*

$$\delta=\big((\beta+\bar\beta z,z),c(\bar\beta,1)\big) \tag{4.2}$$

to G, where $\beta\in\mathbb{D}$ and $c\in\mathbb{C}\setminus\{0\}$. Let ζ,η be the points in $\mathbb{D}$ such that

$$(2\zeta,\zeta^2)=(\beta+\bar\beta\eta,\eta)\in\mathcal{R}\cap F_\beta$$

and let m be the unique automorphism of $\mathbb{D}$ such that

$$m_*((z,c)) = (0, |\delta|_{\text{car}}).$$

For every function $h \in \mathcal{S}(\mathbb{D})$ such that $h(\zeta) = m(\eta)$ the special extension $\tilde{g}$ to G of the function

$$g : \mathcal{R} \cup F_\beta \to \mathbb{D}, \quad (2w, w^2) \mapsto h(w), \quad (\beta + \bar{\beta}w, w) \mapsto m(w) \tag{4.3}$$

for $w \in \mathbb{D}$ is a well-aligned Carathéodory extremal function for δ.

Proof. First observe that there is indeed a unique automorphism m of $\mathbb{D}$ such that $m_*((z,c)) = (0, |\delta|_{\text{car}})$, by the Schwarz-Pick Lemma. Let

$$k(w) = (\beta + \bar{\beta}w, w) \quad \text{for } w \in \mathbb{D},$$

so that $F_\beta = k(\mathbb{D})$ and $k_*((z,c)) = \delta$. By the definition (4.3) of g, $g \circ k = m$.

Consider any function $h \in \mathcal{S}(\mathbb{D})$ such that $h(\zeta) = m(\eta)$. By [2, Lemma 10.5], the function g defined by equations (4.3) is analytic on $\mathcal{R} \cup F_\beta$.

We claim that the special extension $\tilde{g}$ of g to G is a well-aligned Carathéodory extremal function for δ. By [2, Theorem 10.1], $\tilde{g} \in \mathbb{D}(G)$. Moreover

$$\begin{aligned}(\tilde{g})_*(\delta) &= (\tilde{g})_* \circ k_*((z,c)) \\ &= (\tilde{g} \circ k)_*((z,c)) \\ &= (g \circ k)_*((z,c)) \\ &= m_*((z,c)) \\ &= (0, |\delta|_{\text{car}})\end{aligned}$$

as required. Thus the Poincaré metric of $(\tilde{g})_*(\delta)$ on $T\mathbb{D}$ is

$$|(\tilde{g})_*(\delta)| = |(0, |\delta|_{\text{car}})| = |\delta|_{\text{car}}.$$

Therefore $(\tilde{g})_*$ is a well aligned Carathéodory extremal function for δ. □

Clearly the map $g \mapsto \tilde{g}$ is injective, and so this procedure yields a large class of Carathéodory extremals for δ, parametrized by the Schur class.

Remark 4.2. In the converse direction, if φ is any well-aligned Carathéodory extremal for δ, then φ is a norm-preserving extension of its restriction to $\mathcal{R} \cup F_\beta$, which is a function of the type (4.3). Thus the class of all well-aligned Carathéodory extremal functions for δ is given by the set of norm-preserving analytic extensions to G of g in equation (4.3), as h ranges over functions in the Schur class taking the value $m(\eta)$ at

ζ. Typically there will be many such extensions of g, as can be seen from the proof of [2, Theorem 10.1]. An extension is obtained as the Cayley transform of a function defined by a Herglotz-type integral with respect to a probability measure μ on $\mathbb{T}^2$. In the proof of [2, Lemma 10.8], μ is chosen to be the product of two measures $\mu_{\mathcal{R}}$ and $\mu_{\mathcal{F}}$ on $\mathbb{T}$; examination of the proof shows that one can equally well choose any measure μ on $\mathbb{T}^2$ such that

$$\mu(A \times \mathbb{T}) = \mu_{\mathcal{R}}(A), \quad \mu(\mathbb{T} \times A) = \mu_{\mathcal{F}}(A) \quad \text{for all Borel sets } A \text{ in } \mathbb{T}.$$

Thus each choice of $h \in \mathcal{S}(\mathbb{D})$ satisfying $h(\zeta) = m(\eta)$ can be expected to give rise to many well-aligned Carathéodory extremals for δ.

5. Purely balanced tangents

In this section we find a large class of Carathéodory extremals for purely balanced tangents in G by exploiting an embedding of G into the bidisc.

Lemma 5.1. *Let*

$$\Phi = (\Phi_{\omega_1}, \Phi_{\omega_2}) : G \to \mathbb{D}^2$$

where ω_1, ω_2 are distinct points in $\mathbb{T}$. Then Φ is an injective map from G to $\mathbb{D}^2$.

Proof. Suppose Φ is not injective. Then there exist distinct points (s^1, s^2), $(t^1, t^2) \in G$ such that $\Phi_{\omega_j}(s^1, s^2) = \Phi_{\omega_j}(t^1, t^2)$ for $j = 1, 2$. On expanding and simplifying this relation we deduce that

$$s^1 - t^1 - 2\omega_j(s^2 - t^2) - \omega_j^2(s^1 t^2 - t^1 s^2) = 0.$$

A little manipulation demonstrates that both (s^1, s^2) and (t^1, t^2) lie on the complex line

$$\ell \stackrel{\text{def}}{=} \{(s^1, s^2) \in \mathbb{C}^2 : (\omega_1 + \omega_2)s^1 - 2\omega_1\omega_2 s^2 = 2\}.$$

However, ℓ does not meet G. For suppose that $(s^1, s^2) \in \ell \cap G$. Then there exists $\beta \in \mathbb{D}$ such that

$$\begin{aligned} s^1 &= \beta + \bar{\beta}s^2, \\ 2\omega_1\omega_2 s^2 &= (\omega_1 + \omega_2)s^1 - 2 = (\omega_1 + \omega_2)(\beta + \bar{\beta}s^2) - 2. \end{aligned}$$

On solving the last equation for s^2 we find that

$$s^2 = -\bar{\omega}_1\bar{\omega}_2 \frac{2 - (\omega_1 + \omega_2)\beta}{2 - (\bar{\omega}_1 + \bar{\omega}_2)\bar{\beta}},$$

whence $|s^2| = 1$, contrary to the hypothesis that $(s^1, s^2) \in G$. Hence Φ is injective on G. $\square$

Remark 5.2. Φ has an analytic extension to the set $\Gamma\backslash\{(2\bar\omega_1,\bar\omega_1^2),(2\bar\omega_2,\bar\omega_2^2)\}$, where Γ is the closure of G in $\mathbb{C}^2$. However this extension is *not* injective: it takes the constant value $(-\bar\omega_2,-\bar\omega_1)$ on a curve lying in ∂G.

Theorem 5.3. *Let $\delta=(\lambda,v)$ be a purely balanced tangent to G and let Φ_ω solve $\operatorname{Car}\delta$ for the two distinct points $\omega_1,\omega_2\in\mathbb{T}$. Let m_j be the automorphism of $\mathbb{D}$ such that $m_j\circ\Phi_{\omega_j}$ is well aligned at δ for $j=1,2$ and let*

$$\Phi=(\Phi^1,\Phi^2)=(m_1\circ\Phi_{\omega_1},m_2\circ\Phi_{\omega_2}):G\to\mathbb{D}^2. \tag{5.1}$$

For every $t\in[0,1]$ and every function Θ in the Schur class of the bidisc the function

$$F=t\Phi^1+(1-t)\Phi^2+ \\ t(1-t)(\Phi^1-\Phi^2)^2\frac{\Theta\circ\Phi}{1-[(1-t)\Phi^1+t\Phi^2]\Theta\circ\Phi} \tag{5.2}$$

is a well-aligned Carathéodory extremal function for δ.

Proof. By Lemma 5.1, Φ maps G injectively into $\mathbb{D}^2$. By choice of m_j,

$$(m_j\circ\Phi_{\omega_j})_*(\delta)=(0,|\delta|_{\mathrm{car}}).$$

Hence

$$\Phi_*(\delta)=((0,0),|\delta|_{\mathrm{car}}(1,1))\,,$$

which is tangent to the diagonal $\{(w,w):w\in\mathbb{D}\}$ of the bidisc. Since the diagonal is a complex geodesic in $\mathbb{D}^2$, we have

$$|\Phi_*(\delta)|_{\mathrm{car}}=(0,|\delta|_{\mathrm{car}}).$$

As in Section 3, we appeal to [3, Subsection 11.6] to assert that, for every $t\in[0,1]$ and every function Θ in the Schur class of the bidisc, the function $f\in\mathbb{C}(\mathbb{D}^2)$ given by

$$f(\lambda)=t\lambda^1+(1-t)\lambda^2+t(1-t)^2\frac{\Theta(\lambda)}{1-[(1-t)\lambda^1+t\lambda^2]\Theta(\lambda)} \tag{5.3}$$

solves $\operatorname{Car}(\Phi_*(\delta))$. For every such f the function $F\stackrel{\text{def}}{=}f\circ\Phi:G\to\mathbb{D}$ satisfies

$$F_*(\delta)=(f\circ\Phi)_*(\delta)=f_*(\Phi_*(\delta))=(0,|\delta|_{\mathrm{car}}).$$

Thus F is a well-aligned Carathéodory extremal for δ. On writing out F using equation (5.3) we obtain equation (5.2). □

Remark 5.4. The range of Φ is a subset of $\mathbb{D}^2$ containing $(0,0)$ and is necessarily nonconvex, by virtue of a result of Costara [8] to the effect that G is not isomorphic to any convex domain. $\Phi(G)$ is open in $\mathbb{D}^2$, since the Jacobian determinant of $(\Phi_{\omega_1}, \Phi_{\omega_2})$ at (s^1, s^2) is

$$\frac{4(\omega_1 - \omega_2)(1 - \omega_1\omega_2 s^2)}{(2 - \omega_1 s^1)^2(2 - \omega_2 s^1)^2}$$

which has no zero in G. Carathéodory extremals F given by equation (5.3) have the property that the map $F \circ \Phi^{-1}$ on $\Phi(G)$ extends analytically to a map in $\mathbb{D}(\mathbb{D}^2)$. There may be other Carathéodory extremals φ for δ for which $\varphi \circ \Phi^{-1}$ does not so extend. Accordingly we do not claim that the Carathéodory extremals described in Theorem 5.3 constitute all extremals for a purely balanced tangent.

6. Relation to a result of L. Kosiński and W. Zwonek

Our main result in Section 2, on the essential uniqueness of solutions of $\operatorname{Car}\delta$ for purely unbalanced and exceptional tangents, can be deduced from [15, Theorem 5.3] and some known facts about the geometry of G. However, the terminology and methods of Kosiński and Zwonek are quite different from ours, and we feel it is worth explaining their statement in our terminology.

Kosiński and Zwonek speak of left inverses of complex geodesics where we speak of Carathéodory extremal functions for nondegenerate tangents. These are essentially equivalent notions. By a *complex geodesic* in G they mean a holomorphic map from $\mathbb{D}$ to G which has a holomorphic left inverse. Two complex geodesics h and k are *equivalent* if there is an automorphism m of $\mathbb{D}$ such that $h = k \circ m$, or, what is the same, if $h(\mathbb{D}) = k(\mathbb{D})$. It is known (for example [4, Theorem A.10]) that, for every nondegenerate tangent δ to G, there is a *unique* complex geodesic k of G up to equivalence such that δ is tangent to $k(\mathbb{D})$. A function $\varphi \in \mathbb{D}(G)$ solves $\operatorname{Car}\delta$ if and only if $\varphi \circ k$ is an automorphism of $\mathbb{D}$. Hence, for any complex geodesic k and any nondegenerate tangent δ to $k(\mathbb{D})$, to say that k has a unique left inverse up to equivalence is the same as to say that $\operatorname{Car}\delta$ has an essentially unique solution.

Kosiński and Zwonek also use a different classification of types of complex geodesics (or equivalently tangent vectors) in G, taken from [16]. There it is shown that every complex geodesic k in G, up to composition with automorphisms of $\mathbb{D}$ on the right and of G on the left, is of one of the following types.

(1)
$$k(z) = (B(\sqrt{z}) + B(-\sqrt{z}), B(\sqrt{z})B(-\sqrt{z})$$

where B is a non-constant Blaschke product of degree 1 or 2 satisfying $B(0) = 0$;

(2)

$$k(z) = (z + m(z), zm(z))$$

where m is an automorphism of $\mathbb{D}$ having no fixed point in $\mathbb{D}$.

These types correspond to our terminology from [2] (or from Section 1) in the following way. Recall that an automorphism of $\mathbb{D}$ is either the identity, elliptic, parabolic or hyperbolic, meaning that the set $\{z \in \mathbb{D}^- : m(z) = z\}$ consists of either all of $\mathbb{D}^-$, a single point of $\mathbb{D}$, a single point of $\mathbb{T}$ or two points in $\mathbb{T}$.

(1a) If B has degree 1, so that $B(z) = cz$ for some $c \in \mathbb{T}$ then, up to equivalence, $k(z) = (0, -c^2 z)$. These we call the *flat geodesics*. The general tangents to flat geodesics are the flat tangents described in Section 1, that is $\delta = \big((\beta + \bar\beta z, z), c(\bar\beta, 1)\big)$ for some $\beta \in \mathbb{D}$, $z \in \mathbb{D}$ and nonzero $c \in \mathbb{C}$.

(1b) If $B(z) = cz^2$ for some $c \in \mathbb{T}$ then $k(z) = (2cz, c^2 z^2)$. Thus $k(\mathbb{D})$ is the royal variety $\mathcal{R}$, and the tangents to $k(\mathbb{D})$ are the royal tangents.

(1c) If B has degree 2 but is not of the form (1b), say $B(z) = cz(z - \alpha)/(1 - \bar\alpha z)$ where $c \in \mathbb{T}$ and $\alpha \in \mathbb{D} \setminus \{0\}$, then

$$k(z) = \frac{(2c(1 - |\alpha|^2)z, c^2 z(z - \alpha^2))}{1 - \bar\alpha^2 z}.$$

Here $k(\mathbb{D})$ is not $\mathcal{R}$ but it meets $\mathcal{R}$ (at the point $(0, 0)$). It follows that $k(\mathbb{D})$ is a purely unbalanced geodesic and the tangents to $k(\mathbb{D})$ are the purely unbalanced tangents.

(2a) If m is a hyperbolic automorphism of $\mathbb{D}$ then $k(\mathbb{D})$ is a purely balanced geodesic and its tangents are purely balanced tangents.

(2b) If m is a parabolic automorphism of $\mathbb{D}$ then $k(\mathbb{D})$ is an exceptional geodesic, and its tangents are exceptional tangents.

With this description, Theorem 5.3 of [15] can be paraphrased as stating that a complex geodesic k of G has a unique left inverse (up to equivalence) if and only if k is of one of the forms (1c) or (2b). These are precisely the purely unbalanced and exceptional cases in our terminology, that is, the cases of tangents δ for which there is a unique $\omega \in \mathbb{T}$ such that Φ_ω solves $\operatorname{Car}\delta$, in agreement with our Theorem 2.1.

The authors prove their theorem with the aid of a result of Agler and McCarthy on the uniqueness of solutions of 3-point Nevanlinna-Pick problems on the bidisc [3, Theorem 12.13]. They also use the same example from Subsection 11.6 of [3] which we use for different purposes in Sections 3 and 5.

References

[1] J. Agler, Z. A. Lykova and N. J. Young, A geometric characterization of the symmetrized bidisc, preprint.

[2] J. Agler, Z. A. Lykova and N. J. Young, Geodesics, retracts, and the extension property in the symmetrized bidisc, 106 pages, to appear in *Memoirs of the American Mathematical Society*, arXiv:1603.04030 .

[3] J. Agler and J. E. McCarthy, *Pick interpolation and Hilbert function spaces*, Graduate Studies in Mathematics **44**, American Mathematical Society, Providence, RI, 2002.

[4] J. Agler and N. J. Young, The hyperbolic geometry of the symmetrised bidisc, *J. Geometric Analysis* **14** (2004) 375–403.

[5] J. Agler and N. J. Young, Realization of functions on the symmetrized bidisc, *J. Math. Anal. Applic.* **453**(1) (2017) 227– 240.

[6] H. Bart, I. C. Gohberg and M. A. Kaashoek, *Minimal factorization of matrix and operator functions*, Birkhäuser Verlag, Basel, 1979, 277 pp.

[7] T. Bhattacharyya, S. Pal and S. Shyam Roy, Dilations of Γ-contractions by solving operator equations, *Advances in Mathematics* **230** (2012) 577–606.

[8] C. Costara, The symmetrized bidisc and Lempert's theorem, *Bull. Lond. Math. Soc.* **36** (2004) 656–662.

[9] A. Edigarian and W. Zwonek, Geometry of the symmetrized polydisc, *Arch. Math. (Basel)* **84** (2005) 364–374.

[10] A.E. Frazho, S. ter Horst and M.A. Kaashoek, State space formulas for stable rational matrix solutions of a Leech problem, *Indagationes Math.* **25** (2014) 250–274.

[11] M. Jarnicki and P. Pflug, On automorphisms of the symmetrised bidisc, *Arch. Math. (Basel)* **83** (2004) 264–266.

[12] M. Jarnicki and P. Pflug, *Invariant Distances and Metrics in Complex Analysis*, 2nd Extended Edition, De Gruyter, Berlin, 2013.

[13] M.A. Kaashoek and F. van Schagen, The inverse problem for Ellis-Gohberg orthogonal matrix functions, *Integral Equ. Oper. Theory* **80** (2014) 527–555.

[14] S. Kobayashi, *Hyperbolic Complex Spaces*, Grundlehren der mathematischen Wissenschaften **318**, Springer Verlag 1998.

[15] L. Kosiński and W. Zwonek, Nevanlinna-Pick problem and uniqueness of left inverses in convex domains, symmetrized bidisc and tetrablock *J. Geom. Analysis* **26** (2016) 1863–1890.

[16] P. Pflug and W. Zwonek, Description of all complex geodesics in the symmetrized bidisc, *Bull. London Math. Soc.* **37** (2005) 575–584.

[17] J. Sarkar, Operator theory on symmetrized bidisc, *Indiana Univ. Math. J.* **64** (2015) 847–873.

[18] M. Trybula, Invariant metrics on the symmetrized bidisc, *Complex Variables and Elliptic Equations* **60** (4) (2015) 559–565.

Department of Mathematics, University of California at San Diego, CA 92103, USA

School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.
E-mail address: Zinaida.Lykova@ncl.ac.uk

School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K. *and* School of Mathematics, Leeds University, Leeds LS2 9JT, U.K.
E-mail address: Nicholas.Young@ncl.ac.uk