GEOMETRIC PROPERTIES OF LINEAR FRACTIONAL MAPS

C. C. COWEN, D. E. CROSBY, T. L. HORINE, R. M. ORTIZ ALBINO,
A. E. RICHMAN, Y. C. YEOW, B. S. ZERBE

Abstract. Linear fractional maps in several variables generalize classical linear fractional maps in the complex plane. In this paper, we describe some geometric properties of this class of maps, especially for those linear fractional maps that carry the open unit ball into itself. For those linear fractional maps that take the unit ball into itself, we determine the minimal set containing the open unit ball on which the map is an automorphism which provides a means of classifying these maps. Finally, when φ is a linear fractional map, we describe the linear fractional solutions, f, of Schroeder's functional equation $f \circ \varphi = Lf$.

1. Introduction

Linear fractional maps are basic in the theory of analytic functions in the complex plane and in the study of those functions that map the unit disk into itself. We would expect the analogues of these maps in higher dimensions to play a similar role in the study of analytic functions in $\mathbb{C}^N$ and in the study of those functions that carry the unit ball in $\mathbb{C}^N$ into itself. In particular, the linear fractional maps of the unit disk can be classified, most analytic maps of the unit disk into itself inherit one of these classifications, and the classification informs the understanding of the map, for example, of its iteration or the properties of the composition operators with it as symbol. Based on the classification, the solutions of Schroeder's functional equation can be determined and the more general maps of the disk inherit solutions of Schroeder's functional equation.

In this paper, we wish to begin analogous development for several variable analytic maps of the unit ball into itself by building a classification of the several variable linear fractional maps and by solving Schroeder's functional equation for these linear fractional maps. In [4], Cowen and MacCluer define and study a class of linear fractional maps in several variables, relate them to Kreĭn spaces, and prove basic facts about their composition operators; this paper will use their notation and point of view.

Definition A map φ will be called a *linear fractional map* if

$$\varphi(z) = (Az + B)(\langle z, C\rangle + D)^{-1}$$

where A is an $N \times N$ matrix, B and C are (column) vectors in $\mathbb{C}^N$, and D is a complex number. We will regard z as a column vector also and $\langle \cdot, \cdot \rangle$ denotes the usual Euclidean inner product on $\mathbb{C}^N$.

1991 *Mathematics Subject Classification.* Primary 47B38, 32A30; Secondary 30C45.

Supported by National Science Foundation grants DMS-9500870 and DMS-9983601 as well as research support from Purdue University.

Clearly, the domain of a linear fractional map is the set of z in $\mathbb{C}^N$ for which $\langle z, C\rangle + D \neq 0$. The unit ball $\mathbb{B}_N$ in $\mathbb{C}^N$ is the set $\{z : |z| < 1\}$ and the unit sphere is the set $\{z : |z| = 1\}$. In many cases, we want the domain of φ to include the closed ball. Since $z = -DC/|C|^2$ is a zero of $\langle z, C\rangle + D$, for the domain of φ to include the unit ball requires $|-DC/|C|^2| > 1$ or, equivalently, $|D| > |C|$. Conversely, if $|D| > |C|$, then by the Cauchy-Schwarz inequality we will have $\langle z, C\rangle + D \neq 0$ for z in the closed ball. In particular, D is non-zero for these linear fractional maps.

Identifying a 1×1 matrix with its entry, we occasionally write $\langle z, C\rangle = C^* z$. For example, using this identification we can see that a linear fractional map is constant if (and only if) $A = BC^*/D$. We will usually avoid the case of constant maps.

In order to use tools from the theory of Kreĭn spaces, we will sometimes identify $\mathbb{C}^N$ with equivalence classes of points in $\mathbb{C}^{N+1}$. If $v = (v_1, v_2)$ where v_1 is in $\mathbb{C}^N$ and $v_2 \neq 0$ is in $\mathbb{C}$, identify v with v_1/v_2; in particular, $z \leftrightarrow (z, 1)$. We introduce a Kreĭn space structure on $\mathbb{C}^{N+1}$ by letting $[v, w] = \langle Jv, w\rangle$ where $\langle\cdot,\cdot\rangle$ is the usual (Euclidean) inner product on $\mathbb{C}^{N+1}$ and

$$J = \begin{pmatrix} I & 0 \\ 0 & -1 \end{pmatrix}$$

In this setting, v represents a point of the unit sphere if and only if $|v_1| = |v_2|$ which occurs if and only if $[v, v] = |v_1|^2 - |v_2|^2 = 0$ and v represents a point of the unit ball if and only if $[v, v] < 0$.

Definition If $\varphi(z) = (Az + B)(\langle z, C\rangle + D)^{-1}$ is a linear fractional map, the matrix

$$m_\varphi = \begin{pmatrix} A & B \\ C^* & D \end{pmatrix}$$

will be called a *matrix associated with* φ.

Notice that if φ is a linear fractional map with $\varphi(z) = w$ and v is a point of $\mathbb{C}^{N+1}$ associated with z, then $m_\varphi v$ is associated with the point w and vice versa. If φ_1 and φ_2 are linear fractional maps, direct computation of $\varphi_1 \circ \varphi_2$ and $m_{\varphi_1} m_{\varphi_2}$ shows that $\varphi_1 \circ \varphi_2$ is a linear fractional map with associated matrix $m_{\varphi_1 \circ \varphi_2} = m_{\varphi_1} m_{\varphi_2}$. In particular, if φ has a linear fractional inverse, $m_{\varphi^{-1}} = (m_\varphi)^{-1}$ and if m_φ is invertible, φ has a linear fractional inverse.

Note that, if $g(z) = Az$ for some matrix A, then $m_g = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$, or more generally, $m_g = \begin{pmatrix} \alpha A & 0 \\ 0 & \alpha \end{pmatrix}$ for $\alpha \neq 0$, are matrices associated with g.

2. Classification of Linear Fractional Maps

In this section, we consider linear fractional maps that take the unit ball $\mathbb{B}_N$ into itself. An analytic map of the (open) ball to itself may have fixed points in the ball or it may have no fixed points in the ball. If an analytic map, φ, of the ball to itself has no fixed points in the ball, there is a point ζ with $|\zeta| = 1$ such that the iterates of φ converge to ζ uniformly on compact subsets of the ball. In this case, we will refer to ζ as the *Denjoy-Wolff point* of φ. Of course, if φ is continuous on the closed unit ball, such as a linear fractional map that takes the ball into itself, then the Denjoy-Wolff point ζ is a fixed point of φ. If φ has a fixed point in the ball to which the iterates converge (as must be the case when $N = 1$ and φ is not an automorphism), we will also call this point the Denjoy-Wolff point of φ.

In one dimension, linear fractional maps can model those analytic self maps of the disk that have non-zero derivative at their Denjoy-Wolff point. In particular, given $\varphi : \mathbb{D} \to \mathbb{D}$, there exists an open set Ω and an analytic map $\sigma : \mathbb{D} \to \Omega$ such that $\sigma \circ \varphi = \Phi \circ \sigma$, where Φ is a linear fractional map of Ω onto Ω [2, p. 63]. If we also insist that Ω be the smallest set containing $\sigma(D)$ for which $\Phi(\Omega) = \Omega$, then the model parameters (σ, Ω, and Φ) are unique up to holomorphic equivalence. This leads to a classification of analytic maps of the disk into itself according to which linear fractional maps can be used to model the map. Of course, at the same time, this classifies linear fractional maps. The classification results in four cases: half plane / dilation, half plane / translation, plane / dilation, and plane / translation. The domain Ω, which we have taken to be either the complex plane or a half-plane, will be referred to as the *characteristic domain for the model*.

In several variables, the ability to model general analytic maps remains an open question, but some progress has been made [3, 1]. We believe the first step in obtaining such a model is to better understand linear fractional maps in dimension two and higher. In this section, we provide a complete classification of linear fractional maps in two dimensions and a partial classification in N dimensions with the goal of providing the foundation of the classification of analytic maps of the ball into itself.

Thus, we will try to find a small number of linear fractional maps Φ and domains Ω in $\mathbb{C}^N$ such that, given a linear fractional map φ of the ball into itself, we can find one of these Φ and Ω and an open map σ of $\mathbb{B}_N$ into Ω so that

$$\sigma \circ \varphi = \Phi \circ \sigma \tag{1}$$

where $\Phi(\Omega) = \Omega$ and Ω is the smallest set containing $\sigma(\mathbb{B}_N)$ that is invariant for Φ. It is the case, since we are assuming φ is linear fractional, that it is sufficient to take σ to be invertible and a linear fractional map as well. For dimension two, it will turn out that we can choose the characteristic domain Ω to be one of $\mathbb{C}^2$, the half-space $\{(z_1, z_2) : \operatorname{Re} z_1 > 0\}$, or the Siegel half-space $\{(z_1, z_2) : \operatorname{Re} z_1 > |z_2|^2\}$ (which is equivalent to the ball $\mathbb{B}_2$).

Notice that Φ is an automorphism of Ω, that is, Φ is an invertible linear fractional map and $\Phi^{-1}(\Omega) = \Omega$ also. Since $\sigma(\mathbb{B}_N) \subset \Omega$, we get $\Phi^{-1}(\sigma(\mathbb{B}_N)) \subset \Phi^{-1}(\Omega) = \Omega$. Continuing, we get $\Phi^{-n}(\sigma(\mathbb{B}_N)) \subset \Omega$ for every positive integer n. To say that Ω is the smallest domain that satisfies the conditions associated with Equation (1) means that $\Omega = \cup_{n=1}^{\infty} \Phi^{-n}(\sigma(\mathbb{B}_N))$. In particular, for every point w in Ω, there is n so that $\Phi_n(w)$ is in $\sigma(\mathbb{B}_N)$.

Suppose φ is a linear fractional map on $\mathbb{B}_N$ that satisfies $\sigma \circ \varphi = \Phi \circ \sigma$. If η is an automorphism of the ball and $\psi = \eta^{-1} \circ \varphi \circ \eta$, let $\tau = \sigma \circ \eta$. Then

$$\tau \circ \psi = \sigma \circ \eta \circ \eta^{-1} \circ \varphi \circ \eta = \sigma \circ \varphi \circ \eta = \Phi \circ \sigma \circ \eta = \Phi \circ \tau$$

and $\tau(\mathbb{B}_N) = \sigma(\eta(\mathbb{B}_N)) = \sigma(\mathbb{B}_N)$. This means that the data Φ and Ω for the solution of Equation (1) for φ are the same for any ψ that is equivalent to φ under conjugation as above. Moreover, if p is a fixed point of φ, then $q = \eta^{-1}(p)$ is a fixed point of ψ:

$$\psi(q) = \eta^{-1}(\varphi(\eta(\eta^{-1}(p)))) = \eta^{-1}(\varphi(p)) = \eta^{-1}(p) = q$$

In particular, we can normalize the linear fractional maps that we consider with no loss of generality and no change in the classification data. In general, if a linear fractional map has an attractive fixed point in the ball, we will use an automorphism to change the attractive fixed point to be 0. If a linear fractional map has an attractive fixed point ζ on the unit sphere, we will use a rotation to move it to e_1, the "east pole" of the unit ball.

There are two special cases, somewhat degenerate, of linear fractional maps of the ball into itself that will not be considered in this paper. First, we will exclude linear fractional maps that are not invertible as maps of $\mathbb{C}^N$ onto itself; these map the ball into a lower dimensional affine set. We will also exclude linear fractional maps of the ball that do not have a Denjoy-Wolff point; for such maps there is an affine subset of dimension one or more, having non-trivial intersection with the ball, on which the map acts as a generalized rotation.

The easiest case occurs when φ has an attractive fixed point in the interior of the ball, which we can assume to be the origin. In that case, there exists a linear fractional map σ such that $\sigma \circ \varphi = \varphi'(0)\sigma$ [3, Example 3]. That is, Φ in Equation (1) is multiplication by $\varphi'(0)$. Since σ maps $\mathbb{B}_N$ onto a neighborhood of the origin, and the non-degeneracy conditions above mean $\varphi'(0)$ is invertible and all its eigenvalues are less than one in modulus, we see that Ω must be $\mathbb{C}^N$. This case will be referred to as the whole space / dilation case.

For the remainder of this section, we will assume φ has no fixed points in the open ball and that e_1 is the Denjoy-Wolff point of φ, as any boundary fixed point can be moved to e_1 with the use of a generalized rotation and our map could be replaced by the equivalent map obtained by conjugating by this generalized rotation.

To do the full classification, we will make heavy use of the Jordan Canonical Form of m_φ. In particular, when φ is a linear fractional map, m_φ can be factored as

$$m_\varphi = S\Lambda S^{-1} \tag{2}$$

where the columns of S are (generalized) eigenvectors of m_φ and Λ is in Jordan Canonical Form, chosen so that its off-diagonal elements are ones on the subdiagonal. If φ is a linear fractional map of the ball into itself and ψ is an automorphism of the ball, then $\psi\varphi\psi^{-1}$ is another linear fractional map of the ball into itself that is equivalent to φ. Moreover, we see that

$$m_\psi m_\varphi m_\psi^{-1} = m_\psi S\Lambda S^{-1} m_\psi^{-1} = (m_\psi S)\Lambda(m_\psi S)^{-1} \tag{3}$$

so that both φ and $\psi\varphi\psi^{-1}$ yield the same Jordan canonical form matrix.

Recalling from [4, Thm. 9] that the eigenvectors of m_φ correspond to fixed points of φ, we see the case that Λ is diagonal corresponds to the linear fractional map φ having $N+1$ distinct fixed points. This theorem also says that the affine set containing several fixed points of φ is fixed as a set by φ. For maps of the ball into itself, the character of the fixed points different from the Denjoy-Wolff point change depending on whether the line joining the fixed point and the Denjoy-Wolff point is tangent to the ball or not, because if it is not, this line, which is fixed as a set by φ, intersects the interior of the ball. With this in mind, as a definition, if p is a point of the unit sphere, we call a point q in $\mathbb{C}^N$ *p-tangential* if q lies in the hyperspace tangent to the ball at the point p. If the Denjoy-Wolff point of φ is e_1, it is not difficult to show that for all but one column of S, the associated point in

$\mathbb{C}^N$ must be e_1-tangential [7, page 891]. The points in $\mathbb{C}^N$ that are e_1-tangential are precisely those whose first coordinate is 1 and the associated vectors in $\mathbb{C}^{N+1}$ have their first and last coordinates equal. The vectors in $\mathbb{C}^{N+1}$ whose first and last coordinates are zero correspond to the points in the hyperplane at infinity that are e_1-tangential.

In order to proceed with the classification, we first normalize the map by replacing it with a conjugate linear fractional map whose fixed points are more convenient. Equation (3) shows that this is the same as giving S a simpler form. Since we have already assumed the Denjoy-Wolff point of our map is e_1, we will move all other e_1-tangential fixed points off to e_1-tangential points at infinity while keeping e_1 fixed. We will then move the remaining fixed point to an infinite point that is not in the e_1-tangential hyperspace. If some columns of S are generalized eigenvectors which do not correspond to actual fixed points, we will still perform the equivalent algebraic manipulations.

In order to do this normalization, we must be familiar with two particular classes of automorphisms of $\mathbb{B}_N$. The first type, ψ_b, has exactly one fixed point, and corresponds to a Heisenberg translation when considered on the Siegel half-space. That is, $\psi_b = \Psi^{-1} \circ h_b \circ \Psi$ where $\Psi(z) = (z + e_1)/(-z_1 + 1)$, b is on the boundary of the Siegel half-space $\{z : \mathrm{Re}(z_1) > |z_2|^2 + \cdots + |z_N|^2\}$, and h_b is the Heisenberg translation $h_b(z) = Az + b$ where

$$A = \begin{pmatrix} 1 & 2b_2 & \cdots & 2b_N \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

The map Ψ is a biholomorphic map that takes the ball onto the Siegel half-space, $-e_1$ to 0 on the boundary of the Siegel half-space, and e_1 to the infinite point in the direction of e_1, which we sometimes write as $e_{1,\infty}$. For each b, the Heisenberg translation h_b is an automorphism of the Siegel half-space. Moreover, noting that the matrix for h_b is

$$m_{h_b} = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$$

we see that h_b fixes $e_{1,\infty}$ (which corresponds to the vector $(e_1, 0)$ in $\mathbb{C}^{N+1}$), maps the hyperplane at infinity (which corresponds to the hyperplane $\{(v, 0) : v \in \mathbb{C}^N\}$ in $\mathbb{C}^{N+1}$) onto itself since $m_{h_b}(v, 0) = (Av, 0)$, and maps finite points to finite points.

In the equation $m_\varphi = S\Lambda S^{-1}$, the columns of S correspond to the Denjoy-Wolff point e_1, $N-1$ points that are e_1-tangential, and one point that is not e_1-tangential. We want to choose $\psi = \psi_b$ in Equation (3) so that the columns of $m_\psi S$ correspond to e_1, $N-1$ points at infinity that are e_1-tangential, and one point that is not e_1-tangential. A vector corresponds to an e_1-tangential point at infinity exactly when its first and last entries are zero, and each of these corresponds to a fixed point of the biholomorphic map Ψ.

Components $b_2, b_3, \cdots, b_N$ of b are entries of A and we can choose $b_1 = |b_2|^2 + \cdots + |b_N|^2$ to get a point on the boundary of the Siegel half-space. Since $N-1$ of the columns of S correspond to e_1-tangential points, we can find b so that m_{h_b} moves the vectors corresponding to their images under Ψ to vectors whose first and last components are zero. This corresponds to moving the e_1-tangential points of interest to e_1-tangential points at infinity. Replacing S by $m_{\psi_b}S$ corresponds to replacing φ by an equivalent linear fractional map whose fixed points (and their generalizations) are, with a single exception, at e_1 and e_1-tangential points at infinity.

The second type, τ_k, is chosen to fix e_1 as well as the above mentioned hyperspace hyperspace of e_1-tangential points at infinity. To be specific, τ_k is the linear fractional map with associated matrix (in block form)

$$\begin{pmatrix} 1+|k|^2 & 0 & |k|^2-1 \\ 0 & 2kI_{N-1} & 0 \\ |k|^2-1 & 0 & 1+|k|^2 \end{pmatrix}$$

where I_{N-1} is the $N-1$ dimensional identity matrix. With an appropriate choice of k, the column of $m_{\psi_b}S$ corresponding to the point that is not e_1-tangential can be transformed so that it has a 0 as its bottom entry as long as this point was not a second fixed point on the boundary of $\mathbb{B}_N$. This corresponds to moving that point to some infinite point. Incidentally, τ_k can also be used, in conjunction with two rotations to move a second boundary point to $-e_1$ which will later be used for convenience.

Thus, if a linear fractional map of the ball into itself has no fixed point in the open ball and has a single fixed point on the unit sphere, we can replace it with an equivalent linear fractional map of the ball into itself that has its Denjoy-Wolff point at e_1 and its other fixed points (and generalizations) at infinity.

While subjectively different, the two boundary fixed point case is in fact equivalent to the one boundary fixed point case in a neighborhood of the Denjoy-Wolff point. The map η with

$$m_\eta = \begin{pmatrix} 2 & 0 & 0 \\ 0 & I_{N-1} & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

is an automorphism of $\mathbb{C}^N$ together with its points at infinity which fixes e_1, preserves the e_1-tangential hyperplane as a set, and moves $-e_1$ to the infinity associated to $(1,0,0)$. Equivalently, $\eta(\{z : -1 < \operatorname{Re} z_1 < 1\}) = \{z : \operatorname{Re} z_1 < 1\}$ demonstrating the equivalence of the strip (the natural characteristic domain for a non-automorphic linear fractional map fixing $\pm e_1$) and the half-space result given below.

We can therefore say that every linear fractional map φ that has a boundary fixed point is equivalent to one where the S of Equation (2) is of the block form

$$S = \begin{pmatrix} a & 0 & 1 \\ B & C & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where B is $(N-1)\times 1$ and C is $(N-1)\times(N-1)$ unless e_1 and the column $(a, B, 0)$ are both associated with the same Jordan Block. In that case, the columns of S may need to be permuted once more. With one more automorphism, C can be made to be triangular, but this fact is not essential to the following analysis.

We now simply assume that m_φ has been factored as in Equation (2) with S having the form given in the previous paragraph. If Λ is diagonalizable, or even if e_1 is a fixed point of multiplicity one, then the resulting φ must actually be affine. Additionally, after conjugating by translation by e_1 to move the attractive fixed point to the origin, the resulting version of φ can be seen to be multiplication by a matrix with norm less than 1.

In one variable, when φ is linear fractional and the Denjoy-Wolff point is on $\partial\mathbb{D}$ and has multiplicity one, the characteristic domain for the model is always a half-plane. In two dimensions, the characteristic domain is often a half-space, but can also be a Siegel half-space. These will be referred to as the half-space / dilation and Siegel half-space / dilation cases, respectively.

Theorem 1. *If φ is a linear fractional map which has Denjoy-Wolff point on the boundary of $\mathbb{B}_2$ and which has three distinct fixed points, then the characteristic domain for the model will be a half-space or a Siegel half-space.*

Proof. Since φ has three distinct fixed points, the matrix m_φ is diagonalizable. Using a rotation to move the Denjoy-Wolff point to e_1 and the above mentioned automorphisms (τ_k, ψ_b) if necessary, we can assume without loss of generality that the matrix associated with φ has, for some complex number β, the form

$$m_\varphi = \begin{pmatrix} 1 & 0 & 1 \\ \beta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ \beta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}$$

for which the powers, and thus the iterates of φ, are easily calculated. It should also be mentioned that it is easy to determine whether such a φ maps the ball into itself, using the techniques of [7]. Combining the discussion after Theorem 5 with Theorem 1 from [7] and the realization that $0 < \lambda_1 < 1$, we find the necessary condition

$$(1 - \lambda_1)(\lambda_1 - |\lambda_2|^2) - |\lambda_1 - \lambda_2|^2|\beta|^2 \geq 0$$

which immediately implies that $\lambda_1 \geq |\lambda_2|^2$.

In considering the model of Equation (1) in this case, we can temporarily use $\Phi = \varphi$ and $\sigma(z) = z$. Then, to find the characteristic domain of φ, we need to find

$$\Omega = \cup_{n=1}^{\infty} \varphi^{-n}(\sigma(\mathbb{B}_N)) = \cup_{n=1}^{\infty} \varphi^{-n}(\mathbb{B}_N)$$

Or to put it differently, Ω is the set of points for which there is n so that $\varphi_n(w)$ is in $\mathbb{B}_N$. Since $\varphi_n(z_1, z_2) = (\lambda_1^n(z_1 - 1) + 1, \beta(\lambda_1^n - \lambda_2^n)(z_1 - 1) + \lambda_2^n z_2)$, for $\lambda_1 > |\lambda_2|^2$, we find that this is inside the ball for sufficiently large n if and only if $\operatorname{Re} z_1 < 1$, a half-space. When $\lambda_1 = |\lambda_2|^2$, we must have $\beta = 0$ in order for φ to map the ball into itself. It is then easy to calculate that $\varphi_n(z_1, z_2)$ will be in the ball for sufficiently large n if and only if $\operatorname{Re} z_1 < 2 - |z_2|^2$, a Siegel half-space.

For the purposes of the more general model, we normalize both the half-space and the Siegel half-space to $\operatorname{Re} z_1 > 0$ and $\operatorname{Re} z_1 > |z_2|^2$, respectively, so that Φ becomes multiplication by a matrix of norm less than 1. The final version of σ is $\sigma(z) = e_1 - z$. □

As the number of dimensions increases, the necessary conditions and the explicit iteration become rather unwieldy, though evidence supports the following conjecture.

Conjecture 2. *If φ is a linear fractional map which has Denjoy-Wolff point on the boundary of $\mathbb{B}_N$ for which the corresponding matrix m_φ is diagonalizable, then the characteristic domain for the model will be a set of the form $\{z : \operatorname{Re} z_1 > 0\}$ or $\{z : \operatorname{Re}(z_1) > |z_2|^2 + \ldots + |z_k|^2\}$ where $k \leq N$. When $k = N$, this is a classic Siegel half-space. When $k < N$, the regions are known as Siegel domains of Type I.*

It is apparent that the characteristic domain should be at most quadratically bounded since the ball is defined by a quadratic polynomial. That is, when φ is linear fractional, $\varphi_n(z_1, z_2, \dots, z_N)$ will be of the form $(f_1(z)/g(z), \dots, f_N(z)/g(z))$ where f_k and g are affine maps and therefore, $\varphi_n(z_1, \dots, z_N) \subset \mathbb{B}_N$ if and only if $z_1, \dots, z_N$ satisfy a particular quadratic inequality. Letting n tend toward infinity will not change the type of inequality.

When $N = 3$, using the factorization of m_φ from Equation (2), we can show that S must have the form

$$S = \begin{pmatrix} 1 & 0 & 1 \\ B & C & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

If $\lambda_1 > |\lambda_2|^2$ and $\lambda_1 > |\lambda_3|^2$, it becomes a fairly simple calculation to show that the characteristic domain is a half-space. The term which gives rise to this persists in the many variable case. Moreover, when $\lambda_1 = |\lambda_2|^2 = |\lambda_3|^2$, in order for φ to be a self map of the ball, it must be the case that $B = 0$ and $C = I_2$, analogously to the two variable case. In this situation, the iteration is again trivial, and the characteristic domain becomes the Siegel half-space $\{z : \mathrm{Re}(z_1) > |z_2|^2 + |z_3|^2\}$.

For any number of variables, we can take S to have the form

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & I_{n-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

By then choosing $\lambda_1 = |\lambda_j|^2$ when $j \in J$, we get the result that $\varphi_n(z) \in \mathbb{B}_N$ if and only if $\mathrm{Re}\, z_1 < 2 - \sum_{j \in J} |z_j|^2$. We can then permute the variables, which is justified by the symmetry of the problem in $z_2, \dots, z_N$ and normalize to get a characteristic domain of the desired form.

The behavior of a map near a fixed point of multiplicity greater than one is, of necessity, more complicated. Therefore, the model for linear fractional maps with boundary fixed point with multiplicity greater than 1 will be non-trivial. It should be noted that that these maps are generally analyzed more easily on a Siegel half-space than on the ball. In particular, we will use the map Ψ defined above to move e_1 out to a point at infinity. We then become interested in the automorphisms of $\mathbb{C}^N$ which map $\{z : \mathrm{Re}\, z_1 > |z_2|^2 + \ldots + |z_N|^2\}$ into itself and fix the point $e_{1,\infty} = \Psi(e_1)$ which is the point associated to the vector $(1, 0, \dots, 0)$.

By the use of direct calculation when $N = 2$, we find these automorphisms to include the Heisenberg translations, the translations parallel to the axis of the Siegel half-space (i.e. $z \rightarrowtail z + e_1$), and the contractions acting on the second variable (i.e. $(z_1, z_2) \rightarrowtail (z_1, cz_2)$ for $|c| \le 1$ as well as their products. The accompanying table illustrates the possibilities.

In particular, direct calculation shows that when φ is a linear fractional map which has a triple fixed point at e_1, then φ is equivalent to a product of a Heisenberg translation and, if not an automorphism, a translation parallel to the axis. It is easy to see that except in the automorphism case, any point in $\mathbb{C}^2$ will be mapped into the Siegel half-space upon iteration of this product, so $\mathbb{C}^2$ is the characteristic domain. This case will be called the whole space / Heisenberg translation - translation case. The automorphism case will be referred to as the Siegel half-space / Heisenberg translation case.

When φ has a double fixed point, there are two possibilities. If e_1 and the other fixed point have the same eigenvalues associated to them in the matrix factorization, then the map is equivalent to a simple translation. It is again easy to see that the characteristic domain will be the whole space, so we will refer to this case as the whole space / translation case. Finally, the map may be asymptotic to a translation. On the Siegel half-space, these maps are of the form

$$(z_1, z_2) \to (z_1 + a, mz_2 + b)$$

where $\operatorname{Re} a \geq 0$, $|m| \leq 1$, and $|b|^2 \leq (\operatorname{Re} a)(1 - |m|^2)$. These three conditions are necessary and sufficient to guarantee that this maps the Siegel half-space into itself. The first condition is best viewed as a translation term, the second as a compression. The third condition is there to guarantee that any off-axis translation is dominated by the other terms. Regardless of the value of b, an iterate of this function will map any point in $\mathbb{C}^2$ into the Siegel half-space, so the characteristic domain is again the whole space, so we refer to these functions as belonging to the whole space / asymptotic translation case.

Table of examples in two dimensions

Case	Sample $\varphi(z_1, z_2)$
Whole space / dilation	$\varphi(z) = \left(\frac{z_1}{2}, \frac{z_2}{5}\right)$
Half-space / dilation	$\varphi(z) = \left(\frac{z_1 + 4}{5}, \frac{2z_2}{5}\right)$
Siegel half-space / dilation	$\varphi(z) = \left(\frac{z_1 + 3}{4}, \frac{z_1}{2}\right)$
Whole space / Heis. trans.-translation	$\varphi(z) = \left(\frac{z_2 + 1}{-z_1 + z_2 + 2}, \frac{-z_1 + z_2 + 1}{-z_1 + z_2 + 2}\right)$
Siegel half-space / Heis. trans.	$\varphi(z) = \left(\frac{z_1 + 2z_2 + 1}{-z_1 + 2z_2 + 3}, \frac{-2z_1 + 2z_2 + 2}{-z_1 + 2z_2 + 3}\right)$
Whole space / translation	$\varphi(z) = \left(\frac{z_1 + 1}{-z_1 + 3}, \frac{2z_2}{-z_1 + 3}\right)$
Whole space / asymptotic translation	$\varphi(z) = \left(\frac{z_1 + 1}{-z_1 + 3}, \frac{-z_1 + z_2 + 1}{-z_1 + 3}\right)$

In N dimensions, the same compression, translation, and Heisenberg translation factors can be used to create linear fractional maps with a boundary points fixed with any multiplicity. The characteristic domain is then determined by the various factors. In particular, eigenvalues of modulus 1 in a contraction factor can result in different characteristic domains which are of Siegel type I, analogous to those mentioned in Conjecture 2.

3. Schroeder's Equation

For an arbitrary analytic map φ of the disk $\mathbb{D}$ to itself, fixing 0 and with $\varphi'(0) = \lambda$ satisfying $0 < |\lambda| < 1$, Koenigs [6] in 1884 gave an essentially unique analytic function f in $\mathbb{D}$ solving Schroeder's functional equation

$$f \circ \varphi = \lambda f \tag{4}$$

In the years since then, all analytic solutions of Schroeder's functional equation have been found for analytic maps of the disk into itself.

When φ is an analytic map of $\mathbb{B}_N$ into $\mathbb{B}_N$, we may seek a $\mathbb{C}^N$-valued analytic function f on $\mathbb{B}_N$ solving the several variable Schroeder equation

$$f \circ \varphi = Lf \tag{5}$$

where L is an $N \times N$ matrix. In [4, Theorem 19], Cowen and MacCluer give the following theorem as a solution of Schroeder's functional equation in a special case.

Theorem 3. *Suppose $\varphi : \mathbb{B}_N \to \mathbb{B}_N$ is a linear fractional map with $\varphi(0) = 0$. Then there is an invertible linear fractional map f defined in a neighborhood of 0, with $f \circ \varphi = \varphi'(0)f$. Moreover, if no eigenvalue of $\varphi'(0)$ has modulus 1, then the domain of f includes the unit ball, $\mathbb{B}_N$.*

Our goal in this section is to find, for linear fractional maps φ, all invertible linear fractional solutions f of Schroeder's functional equation in several variables. The following theorem describes how various solutions of Schroeder's equation might be related to other solutions. The remaining results of the section complete the goal of finding linear fractional solutions of Schroeder's functional equation when φ is a linear fractional map.

Theorem 4. *Suppose L_1 is an $N \times N$ matrix and suppose f_1 is an invertible linear fractional solution of $f_1 \circ \varphi = L_1 f_1$. If L_2 is a matrix similar to L_1, say $L_2 = PL_1P^{-1}$ for some invertible $N \times N$ matrix P, then $f_2 = Pf_1$ is a solution of $f_2 \circ \varphi = L_2 f_2$. Conversely, if L_1 has no eigenvalues that are roots of unity, L_2 is similar to L_1, and f_1, f_2 are invertible functions satisfying $f_1 \circ \varphi = L_1 f_1$ and $f_2 \circ \varphi = L_2 f_2$, then there is an invertible matrix P so that $L_2 = PL_1P^{-1}$ and $f_2 = Pf_1$.*

Proof. Let $m_{f_1}, m_{f_2}, m_\varphi, m_{L_1}$, and m_{L_2} be the matrices associated with the linear fractional maps f_1, f_2, φ and multiplication by L_1 and L_2, respectively. Without loss of generality, we choose the constants so that

$$m_{L_1} = \begin{pmatrix} L_1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad m_{f_1} m_\varphi = m_{L_1} m_{f_1} \quad \text{and} \quad m_{f_2} m_\varphi = m_{L_2} m_{f_2}$$

Then $m_\varphi = m_{f_1}^{-1} m_{L_1} m_{f_1} = m_{f_2}^{-1} m_{L_2} m_{f_2}$ so that

$$\left(m_{f_2} m_{f_1}^{-1}\right) m_{L_1} = m_{L_2} \left(m_{f_2} m_{f_1}^{-1}\right) \tag{6}$$

Suppose that in block form

$$\left(m_{f_2} m_{f_1}^{-1}\right) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where A is $N \times N$. We then have that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} L_1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha L_2 & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where $\alpha \neq 0$ and therefore

$$AL_1 = \alpha L_2 A,\ B = \alpha L_2 B,\ CL_1 = \alpha C,\ \text{and } D = \alpha D.$$

Since L_1 and L_2 were assumed to be similar, they have the same set of eigenvalues, say $\mu_1, \mu_2, \dots, \mu_k$. From Equation (6), we have that m_{L_1} is similar to m_{L_2} and therefore have the same eigenvalues. Since

$$m_{L_1} = \begin{pmatrix} L_1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } m_{L_2} = \begin{pmatrix} \alpha L_2 & 0 \\ 0 & \alpha \end{pmatrix}$$

we see that the eigenvalues of m_{L_1} are $\{1, \mu_1, \mu_2, \dots, \mu_k\}$ while the eigenvalues of m_{L_2} are $\{\alpha, \alpha\mu_1, \dots, \alpha\mu_k\}$. Since these sets are the same, and indeed, the same as $\{\alpha^j, \alpha^j\mu_1, \dots, \alpha^j\mu_k\}$ for every integer j, we see that α must be a root of unity. Since no root of unity is an eigenvalue of L_1, no root of unity can be an eigenvalue of L_2. Thus, $\alpha L_2 B = B$ implies that $B = 0$. Similarly, $CL_1 = \alpha C$ implies $L_1^* C^* = \overline{\alpha} C^*$ and no root of unity is an eigenvalue of L_1^* so $C^* = 0$ and $C = 0$. Hence,

$$m_{f_2} m_{f_1}^{-1} = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \text{ and } m_{f_2} = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} m_{f_1}.$$

Thus by taking $P = D^{-1}A$, we have $f_2 = Pf_1$ and since $AL_1 = L_2 A$, we also have $L_2 = PL_1P^{-1}$. □

It is worth examining the reason for the hypothesis that no root of unity is an eigenvalue of L_1. The reason is evident with the following example: Let $f_1(z_1, z_2) = (z_1, z_2)$, let $f_2(z_1, z_2) = (z_2/z_1, 1/z_1)$, and let $\varphi(z_1, z_2) = (\zeta z_1, \zeta^2 z_2)$ where $\zeta^3 = 1$ and $\zeta \neq 1$. Also, consider

$$L_1 = L_2 = \begin{pmatrix} \zeta & 0 \\ 0 & \zeta^2 \end{pmatrix}$$

It can be easily verified that $m_{f_1} m_\varphi = m_{L_1} m_{f_1}$ where a 1 is in the lower right corner of each of these matrices. Using

$$m_{L_2} = \begin{pmatrix} \alpha L_2 & 0 \\ 0 & \alpha \end{pmatrix}$$

we easily obtain that $m_{f_2} m_\varphi = m_{L_2} m_{f_2}$ when $\alpha = \zeta$. It is of course clear that f_2 can not be written as Pf_1 for any constant 2×2 matrix P.

In the theorem below, $[v]$ denotes the subspace spanned by the vector v.

Theorem 5. *Suppose φ is a linear fractional map on $\mathbb{C}^N$. There is an $N \times N$ matrix L and an invertible linear fractional map f so that $f \circ \varphi = Lf$ if and only if there is an eigenvector v for m_φ with non-zero eigenvalue and a subspace $M \subset \mathbb{C}^{N+1}$ such that M is invariant for m_φ and $M \oplus [v] = \mathbb{C}^{N+1}$. In this case, L is similar to the restriction of m_φ to M.*

Proof. ($\Rightarrow$) Suppose φ, f, and L are as in the hypothesis and satisfy $f \circ \varphi = Lf$. Choose the constants so that $m_f m_\varphi = m_L m_f$ for

$$m_L = \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix}$$

Since f is invertible, we can rewrite the above equation as $m_\varphi m_f^{-1} = m_f^{-1} m_L$. Denote the first N columns of m_f^{-1} by $u_1, u_2, \dots, u_N$ and the last column by v. Let $M = \operatorname{span}\{u_1, u_2, \dots, u_N\}$. Since the columns of m_f^{-1} are linearly independent, we have immediately that $M \oplus [v] = \text{span}\{u_1, u_2, \dots, u_N, v\} = \mathbb{C}^{N+1}$. Blocking m_f^{-1} by columns, we get

$$m_\varphi m_f^{-1} \begin{pmatrix} u_1 & u_2 & \cdots & u_N & v \end{pmatrix} = \begin{pmatrix} m_\varphi u_1 & m_\varphi u_2 & \cdots & m_\varphi u_N & m_\varphi v \end{pmatrix}.$$

On the other hand,

$$\begin{aligned} m_f^{-1} m_L &= \begin{pmatrix} u_1 & u_2 & \cdots & u_N & v \end{pmatrix} \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1N} & 0 \\ l_{21} & l_{22} & \cdots & l_{2N} & 0 \\ \vdots & & & \vdots & \vdots \\ l_{N1} & l_{N2} & \cdots & l_{NN} & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^N l_{j1} u_j & \sum_{j=1}^N l_{j2} u_j & \cdots & \sum_{j=1}^N l_{jN} u_j & v \end{pmatrix} \end{aligned}$$

Comparing the two expressions, we see that

$$m_\varphi u_j = \sum_{j=1}^N l_{jk} u_j$$

for each $k = 1, \dots, N$, so M is invariant for m_φ. Also, we see that v is an eigenvector of m_φ with eigenvalue 1.

($\Leftarrow$) Conversely, suppose that v is an eigenvector for m_φ with eigenvalue $\alpha \neq 0$ and M is an invariant subspace for m_φ such that $M \oplus [v] = \mathbb{C}^{N+1}$. Let $\{u_1, u_2, \dots, u_N\}$ be a basis for M so that $\{u_1, u_2, \dots, u_N, v\}$ will be a basis for $\mathbb{C}^{N+1}$. Let S be the matrix whose columns are $u_1, u_2, \dots u_N, v$. The matrix S is invertible and the choice of its columns means that in block form,

$$S^{-1} m_\varphi S = \begin{pmatrix} L_0 & 0 \\ 0 & \alpha \end{pmatrix}$$

for some $N \times N$ matrix L_0. The right hand side matrix is m_L for $L = \alpha^{-1} L_0$ and defining a linear fractional map f by $m_f = S^{-1}$, we see that $m_f m_\varphi = m_L m_f$. That is, f is an invertible linear fractional map and L is an $N \times N$ matrix such that $f \circ \varphi = Lf$.

□

By the basic Jordan block of size l associated with the eigenvalue λ, we mean the $l \times l$ matrix whose diagonal entries are λ and whose subdiagonal entries are 1. The Jordan Canonical Form Theorem says that every matrix is similar to a block diagonal matrix whose diagonal entries are basic Jordan blocks, and that the set of basic Jordan blocks is completely determined by the matrix. If A is a matrix, we say that "A has basic Jordan blocks $J_1, \dots, J_k$" if these are the blocks that occur in the Jordan canonical form for A.

Corollary 6. *Let φ be a linear fractional map on $\mathbb{C}^N$, let m_φ be a matrix representing φ, and suppose m_φ has basic Jordan blocks $J_1, J_2, \dots, J_k$ associated with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$. Let L be an $N \times N$ matrix.*

There is an invertible linear fractional map f satisfying $f \circ \varphi = Lf$ if and only if there is a 1×1 Jordan block J_j with $\lambda_j \neq 0$ for which $\lambda_j L$ has basic Jordan blocks $J_1, \dots, J_{j-1}, J_{j+1} \dots, J_k$.

Proof. The corollary follows immediately from the proof of Theorem 5. Each basic Jordan block of m_φ corresponds to an invariant subspace for m_φ, with the 1×1 blocks corresponding to an eigenspace of m_φ. By taking the eigenvector v to be an eigenvector associated with the 1×1 Jordan block J_j (and putting λ_j in the lower right corner of the corresponding matrix) and the subspace M to be the subspace spanned by the invariant subspaces associated with the basic Jordan blocks $J_1, \ldots, J_{j-1}, J_{j+1} \ldots, J_k$, the proof of Theorem 5 shows that the resulting matrix m_L is

$$m_L = \begin{pmatrix} \lambda_j L & 0 \\ 0 & \lambda_j \end{pmatrix}$$

which is similar to m_φ and has the same basic Jordan blocks. In particular, the matrix $\lambda_j L$ has the blocks of m_φ except J_j. □

Corollary 7. *Let φ be a linear fractional map on $\mathbb{C}^N$ such that m_φ is invertible and diagonalizable and let L be an $N \times N$ matrix.*

If the eigenvalues of m_φ are λ_1, λ_2, $\cdots$, and λ_{N+1}, then there is an invertible linear fractional map f satisfying $f \circ \varphi = Lf$ if and only if L is a diagonalizable matrix whose eigenvalues are $\{\frac{\lambda_1}{\lambda_j}, \frac{\lambda_2}{\lambda_j}, \ldots, \frac{\lambda_{j-1}}{\lambda_j}, \frac{\lambda_{j+1}}{\lambda_j}, \ldots, \frac{\lambda_{N+1}}{\lambda_j}\}$ for some $j = 1, 2, \ldots, N+1$.

Proof. Since m_φ is diagonalizable, the basic Jordan blocks for m_φ are all 1×1. This corollary is a restatement of the previous corollary with the observation that the basic Jordan block J_ℓ is just (λ_ℓ) for each $\ell = 1, 2, \ldots, N+1$. □

The ideas of Theorem 5 have other consequences as well. Suppose m_φ and f are linear fractional maps of $\mathbb{C}^N$ with f invertible, suppose L is an $N \times N$ matrix for which 1 is not an eigenvalue, and suppose they satisfy Schroeder's equation $f \circ \varphi = Lf$. If p is a fixed point of φ that is in the domain of f, then $Lf(p) = f(\varphi(p)) = f(p)$. Since 1 is not an eigenvalue of L, this means $f(p) = 0$. Moreover, since f is invertible, p is the only point of $\mathbb{C}^N$ that f maps to 0. At the level of m_φ, referring to the proof of Theorem 5, the vector 0 in $\mathbb{C}^N$ corresponds to the vector $z_0 = (0, 0, \ldots, 0, 1)$ in $\mathbb{C}^{N+1}$, and v is the vector for which $m_f(v) = z_0$. In other words, the vector v represents the fixed point p of φ.

Because there is only one point that f maps to 0 and f maps every fixed point of φ that is in its domain to 0, all the fixed points of φ besides the fixed point p must fail to be in the domain of f. Moreover, an affine combination of points not in the domain of f will also fail to be in the domain of f, so the affine set that is spanned by the fixed points of φ besides p is in not in the domain of f. Perhaps more clearly, at the level of m_φ, referring to the proof of Theorem 5, vectors not in the domain of f correspond to the vectors w in $\mathbb{C}^{N+1}$ such that $m_f(w)$ is a vector whose last component is 0. The proof of Theorem 5 shows that this is exactly the subspace M. In other words, the vectors not in the domain of f are the vectors of $\mathbb{C}^N$ represented by vectors in M: this will be a hyperplane in $\mathbb{C}^N$, it is the hyperplane $\{z : C^*z + D = 0\}$ if

$$m_\varphi = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

The construction of M shows that this hyperplane includes all the fixed points of φ besides p.

References

[1] F. Bracci and G. Gentili, Solving the Schröder equation at the boundary in several variables, preprint, 2003.

[2] C. C. Cowen and B. D. MacCluer, "Composition Operators on Spaces of Analytic Functions," CRC Press, Boca Raton, 1995.

[3] C. C. Cowen and B. D. MacCluer, Schroeder's equation in several variables, *Taiwanese J. Math.* **7**(2003), 129–154.

[4] C. C. Cowen and B. D. MacCluer, Linear fractional maps of the ball and their composition operators, *Acta Sci. Math. (Szeged)* **66**(2000), 351–376.

[5] L. A. Harris, Linear fractional transformations of circular domains in operator spaces, *Indiana Univ. Math. J.* **41**(1992), 125–147.

[6] G. Koenigs, Recherches sur les intégrales de certaines équations functionelles, *Annales Ecole Normale Superior (Sér. 3)* **1**(1884), supplément, 3–41.

[7] A. E. Richman, The range of linear fractional maps on the unit ball, *Proc. Amer. Math. Soc.* **131**(2003), 889–895.

[8] W. Rudin, "Function Theory in the Unit Ball of C^n," Grundlehren der math. Wiss. **241**, Springer–Verlag, New York, 1980.

[9] E. M. Stein, "Boundary Behavior of Holomorphic Functions of Several Complex Variables," Princeton University Press, Princeton, 1972.

Purdue University, West Lafayette, Indiana 47907
E-mail address: cowen@math.purdue.edu

Bucknell University, Lewisburg, Pennsylvania 17837
E-mail address: arichman@bucknell.edu