PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 131, Number 3, Pages 889–895
S 0002-9939(02)06596-6
Article electronically published on July 17, 2002

THE RANGE OF LINEAR FRACTIONAL MAPS ON THE UNIT BALL

ALEXANDER E. RICHMAN

(Communicated by Joseph A. Ball)

Abstract. In 1996, C. Cowen and B. MacCluer studied a class of maps on $\mathbb{C}^N$ that they called linear fractional maps. Using the tools of Kreĭn spaces, it can be shown that a linear fractional map is a self-map of the ball if and only if an associated matrix is a multiple of a Kreĭn contraction. In this paper, we extend this result by specifying this multiple in terms of eigenvalues and eigenvectors of this matrix, creating an easily verified condition in almost all cases. In the remaining cases, the best possible results depending on fixed point and boundary behavior are given.

1. Introduction

In one dimension, the behavior of analytic self-maps of the unit disk is in many ways characterized by the location of its fixed points and the derivative at those fixed points. In turn, the various possibilities can be modeled by linear fractional maps [5]. This is particularly important in the study of composition operators. It is hoped that analytic self-maps of the unit ball can be characterized by a class of more easily understood functions. In [7], C. Cowen and B. MacCluer proposed, using the following N-dimensional analogue of linear fractional maps, to study composition operators:

Let A be an $N \times N$ matrix, let B and C be $N \times 1$ column vectors, and let D be a complex number. Then the function $\varphi : \mathbb{C}^N \to \mathbb{C}^N$ defined by

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

is called a *linear fractional map*. These maps have been studied in more generality by many others [12], [15], [10], [13].

In order to study these maps, Cowen and MacCluer used a Kreĭn space structure. In particular, they identified points in $\mathbb{C}^N$ with an equivalence class of points in $\mathbb{C}^{N+1}$. Two points, $v, w \in \mathbb{C}^{N+1}$, are considered to be equivalent if $v = cw$ for some $c \in \mathbb{C}$. The relevant Kreĭn inner product is given by $[v, w] = \langle Jv, w\rangle$ where $\langle \cdot, \cdot \rangle$ is the usual Euclidean inner product on $\mathbb{C}^{N+1}$ and

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

Received by the editors September 12, 2001 and, in revised form, October 19, 2001.
2000 *Mathematics Subject Classification.* Primary 32A10, 32A40, 47B50.
Key words and phrases. Linear fractional maps, unit ball, Kreĭn space.

They further introduce a *matrix associated with* φ given by

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

which is unique up to multiplication by a constant. The value of this notation is apparent when noting that for any $z \in \mathbb{C}^N$, $m_\varphi \binom{z}{1}$ is in the same equivalence class as $\binom{\varphi(z)}{1}$. In particular, this forces the eigenvectors of m_φ to represent the fixed points of φ.

A major difficulty in the study of analytic functions on the unit ball, $\mathbb{B}_N$, is the problem of determining whether such a function maps the ball into itself. This is the first critical question in the study of composition operators in several variables. A linear fractional map $\varphi : \mathbb{C}^N \to \mathbb{C}^N$ maps the ball into itself if and only if m_φ is a multiple of a Kreĭn contraction; that is, $t^2[m_\varphi v, m_\varphi v] \leq [v, v]$ for some $t > 0$. This was shown in a general Kreĭn space setting by [12] and [15]. In the current context, it appeared in [7, Theorem 13]. However, they give no guidance as to which multiple should be selected. This is particularly important since exactly one positive choice of t will work for most linear fractional self-maps of the ball. In [2], C. Bisi and F. Bracci give a geometric characterization of linear fractional self-maps of the ball, but their results do not yield a practical manner by which one can determine whether a given linear fractional map does in fact take $\mathbb{B}_N$ into itself. In 1996, D. Crosby, an undergraduate acting under the direction of Cowen, obtained the first definitive results of this form, which were limited to φ with real valued coefficients and a boundary fixed point in the case $N = 2$ [8].

It is well known that any analytic self-map of $\mathbb{B}_N$ must fix at least one point in the closed unit ball. In this paper, we consider two substantially different cases. First, if φ has a fixed point on $\partial\mathbb{B}_N$, then the positive multiple t for which tm_φ is a Kreĭn contraction is uniquely determined. Specifically, if $d \in \mathbb{C}^{N+1}$ is a representation of a boundary fixed point, λ is the eigenvalue of m_φ corresponding to d, and x is any vector in $\mathbb{C}^{N+1}$ such that $[d, x] \neq 0$, then $t^2 = \operatorname{Re}([d, x])/\operatorname{Re}(\lambda[d, m_\varphi x])$. It is worth noting that the ratio $[d, x]/[d, m_\varphi x]$ is itself an eigenvalue of m_φ.

When φ has no boundary fixed points, the multiplier t is in general not unique. In fact, the best possible information based on eigenvector and eigenvalue information alone is that $1/|\lambda_1| \leq t \leq 1/|\lambda_2|$ where λ_1 is the largest eigenvalue (which automatically corresponds to the interior fixed point) and λ_2 is the next largest eigenvalue. More precise information is available if in addition, one boundary point is mapped to another boundary point. In this case, we again are forced to have $t^2 = \operatorname{Re}([d, x])/\operatorname{Re}(\lambda[d, m_\varphi x])$ where in this case d represents the interior fixed point but x is not arbitrary. Instead, x must be chosen so that x and $m_\varphi x$ both represent boundary points.

2. Boundary fixed points

Throughout this section, we will assume that there is a point $p \in \partial\mathbb{B}_N$ such that $\varphi(p) = p$. For notational simplicity, we will refer to $\binom{p}{1}$ as d so that d is a vector in $\mathbb{C}^{N+1}$ associated with p. The geometric meaning of each distinct fixed point changes depending on whether the line between it and p is tangent to the ball since 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, 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 v is the vector associated with

q, then this condition is satisfied exactly when $[d,v]=0$ because for $v=\binom{kq}{k}$, we have that

$$\begin{aligned}[d,v] &= \overline{k}\,(\langle p,q\rangle - 1)\\ &= \overline{k}\,(\langle p,q\rangle - \langle p,p\rangle)\\ &= \overline{k}\,(\langle p,q-p\rangle)\end{aligned}$$

which is 0 precisely when the vector from the point associated with v to p is perpendicular to the normal vector at p. In [7], [9] and [11], it is shown that whenever φ is a self-map of $\mathbb{B}_N$, the Kreĭn adjoint of m_φ also corresponds to a self-map of $\mathbb{B}_N$. Though not obvious, it follows that m_φ has exactly one generalized eigenvector (up to a constant multiple) corresponding to a point which is not p-tangential. Note that if φ has two boundary fixed points, neither is in the tangent hyperspace associated with the other. It does not matter which one is chosen in the result that follows. This allows us to state our primary result.

Theorem 1. *Let φ be a linear fractional map with fixed point $p\in\partial\mathbb{B}_N$ and associated matrix m_φ. Let λ_1 be the eigenvalue (which can be assumed to be real) of m_φ corresponding to $d=\binom{p}{1}$ and let $x\in\mathbb{C}^{N+1}$ be any vector such that $[d,x]\neq 0$. Then $\varphi(\mathbb{B}_N)\subset\mathbb{B}_N$ if and only if tm_φ is a Kreĭn contraction, where*

$$t=\sqrt{\frac{\operatorname{Re}([d,x])}{\operatorname{Re}(\lambda_1[d,m_\varphi x])}}.$$

Moreover, $\operatorname{Re}([d,m_\varphi x])/\operatorname{Re}([d,x])$ *is the eigenvalue corresponding to the generalized eigenspace containing a generalized eigenvector of m_φ which is not Kreĭn orthogonal to d.*

Proof. It is clear that if φ is not a self-map of the ball, then tm_φ will not be a Kreĭn contraction for any $t>0$. Therefore, it suffices to assume that $\varphi(\mathbb{B}_N)\subset\mathbb{B}_N$. We can then assume that there exists a $t>0$ such that $t^2[m_\varphi v,m_\varphi v]\le[v,v]$ for all $v\in\mathbb{C}^{N+1}$. Choose $x\in\mathbb{C}^{N+1}$ so that $[d,x]\neq 0$, that is, x corresponds to a point which is not p-tangential, and let $y=m_\varphi x$. We then set $v=sx+(1-s)d$ so that $m_\varphi v=sy+(1-s)\lambda_1 d$. Therefore,

$$\begin{aligned}[m_\varphi v,m_\varphi v] &= s^2[y,y]+s(1-s)[y,\lambda_1 d]+s(1-s)[\lambda_1 d,y]+(1-s)^2|\lambda_1|^2[d,d]\\ &= s^2[y,y]+2s(1-s)\operatorname{Re}(\lambda_1[d,y])\end{aligned}$$

since $[d,d]=\langle p,p\rangle-1=0$ because p is on $\partial\mathbb{B}_N$. Similarly,

$$[v,v]=s^2[x,x]+2s(1-s)\operatorname{Re}([d,x]).$$

For sufficiently small values of s, we may ignore the quadratic term and rewrite the inequality $t^2[m_\varphi v,m_\varphi v]\le[v,v]$ to get that

$$t^2 s\operatorname{Re}(\lambda_1[d,y])\le s\operatorname{Re}([d,x]).$$

By choosing s to have the same sign as $\operatorname{Re}(\lambda_1[d,y])$ we get

$$t^2\le\frac{\operatorname{Re}([d,x])}{\operatorname{Re}(\lambda_1[d,y])}$$

and by choosing s to have the opposite sign, we get the reverse inequality. Therefore,

$$t^2=\frac{\operatorname{Re}([d,x])}{\operatorname{Re}(\lambda_1[d,y])}. \tag{2.1}$$

Note that if the right-hand side of this expression were not independent of x or if it were negative, this would be a contradiction with the assumption that there exists a $t > 0$ such that $t^2[m_\varphi v, m_\varphi v] \leq [v, v]$ for all $v \in \mathbb{C}^{N+1}$, completing the first part of the proof.

To show that $\operatorname{Re}([d, m_\varphi x])/\operatorname{Re}([d, x])$ is an eigenvalue, we need to use linear algebra to give some structure to m_φ. Specifically, we can express it in the following Jordan canonical form:

$$m_\varphi = E\Lambda E^{-1} \tag{2.2}$$

where d is the first (left-most) column of E and the i-th column of E will be denoted by v_i so that v_i is a generalized eigenvector of m_φ and Λ has the form:

$$\Lambda = \begin{pmatrix} \Lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \Lambda_K \end{pmatrix}, \Lambda_j = \begin{pmatrix} \lambda_j & 1 & 0 & \dots & 0 \\ 0 & \lambda_j & 1 & 0 \dots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \lambda_j & 1 \\ 0 & 0 & \dots & 0 & \lambda_j \end{pmatrix}.$$

Moreover, if v_M is the unique column of E which is not Kreĭn orthogonal to d, then the M-th row of Λ must be the bottom row of a Jordan block, say Λ_m. This also follows from the fact that the Kreĭn adjoint of a self-map of the ball is itself a self-map of the ball [7], [11], [9]. Furthermore, the M-th row of E^{-1} will be a constant times the vector $(p, -1)$. Now, we can rewrite the equation $x = m_\varphi y$ as

$$E^{-1}x = \Lambda E^{-1}y.$$

By looking at the entry in row M of each side of this equation, we get

$$[d, x] = \lambda_m [d, y].$$

Since it is also true that $\operatorname{Re}([d, x])/\operatorname{Re}([d, y])$ is constant and x can be replaced by cx for any $c \in \mathbb{C}, c \neq 0$ in each equation, it follows that

$$\frac{\operatorname{Re}([d, x])}{\operatorname{Re}([d, y])} = \frac{[d, x]}{[d, y]}$$

and that the ratio $\lambda_m/\lambda_1 \in \mathbb{R}$ for any self-map of the ball, which completes the proof. □

In practical terms, this theorem says that if φ has a boundary fixed point, one need only check to see whether a specific matrix is a Kreĭn contraction. Equivalently, as noted in [7], it suffices to check whether the matrix $J - t^2 m_\varphi^* J m_\varphi$ has only non-negative eigenvalues for $t^2 = [d, x]/(\lambda_1 [d, m_\varphi x])$.

We note that this can be written in a more general setting as follows.

Theorem 2. *Let $\mathcal{K}$ be a Kreĭn space with finitely many negative squares and let $M : \mathcal{K} \to \mathcal{K}$ be strict-minus. That is, $[Mx, Mx] < 0$ whenever $[x, x] < 0$. Suppose that there exists $d \in \mathcal{K}$ such that $Md = \lambda_1 d$ and $[d, d] = 0$. If $x \in \mathcal{K}$ is any element such that $[d, x] \neq 0$, then tm_φ is a Kreĭn contraction if and only if*

$$t = \sqrt{\frac{\operatorname{Re}([d, x])}{\operatorname{Re}(\lambda_1 [d, m_\varphi x])}}.$$

The proof is identical to the proof of Theorem 1.

3. No boundary fixed points

Unfortunately, when φ has no boundary fixed points, the situation is somewhat more difficult since generalized eigenvector and eigenvalue information alone is not sufficient to say which multiples of m_φ would be Kreĭn contractions. On the other hand, this multiple in general is not unique, allowing computer assisted guesswork to easily deliver a working multiple if one exists. More on this can be found below, following the proof of Theorem 5.

Theorem 3. *If φ is a linear fractional self-map of the unit ball without boundary fixed points and tm_φ is a Kreĭn contraction for some $t > 0$, then $1/|\lambda_1| \leq t \leq 1/|\lambda_2|$ where λ_1 is the eigenvalue corresponding to the interior fixed point and λ_2 is an eigenvalue with largest absolute value corresponding to an exterior fixed point.*

Proof. We begin the proof with two observations. First, $|\lambda_1| \geq |\lambda_2|$ since the interior fixed point must be attractive. Second, for any vector $v \in \mathbb{C}^{N+1}$, $[v, v] < 0$ if and only if v represents a point inside the ball. Likewise, $[v, v] > 0$ if and only if v is associated to a point outside the ball. Next, we let v_1, v_2 be eigenvectors with eigenvalues λ_1, λ_2, respectively. We therefore have that

$$t^2[m_\varphi v_1, m_\varphi v_1] = t^2|\lambda_1|^2[v_1, v_1] \leq [v_1, v_1]$$

which quickly yields that $t \geq 1/|\lambda_1|$ since $[v_1, v_1] < 0$. Similarly, we get

$$t^2[m_\varphi v_2, m_\varphi v_2] = t^2|\lambda_2|^2[v_2, v_2] \leq [v_2, v_2].$$

However, since $[v_2, v_2] > 0$, this yields that $t \leq 1/|\lambda_2|$. □

To demonstrate why this is the best possible information based on eigenvalue information alone, consider the following example:

Example 4. For $b > 1$, let

$$m_\varphi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & \frac{-4}{b} & 6 \end{pmatrix}.$$

Each of these has eigenvalues 1, 2, and 6 with one interior fixed point and two exterior fixed points. As b approaches 1, this approaches the boundary fixed point case. As a result, $(1/\sqrt{12})m_\varphi$ is a Kreĭn contraction but for any other t with $1/6 \leq t \leq 1/2$, there is a $b > 1$ such that tm_φ is not a Kreĭn contraction. To complete this example, we consider

$$m_\varphi = \begin{pmatrix} 2 & -\frac{16}{5} & 0 \\ 0 & 1 & 0 \\ \frac{-4}{3} & -\frac{12}{5} & 6 \end{pmatrix}.$$

We see that this also has an interior fixed point and two exterior fixed points with eigenvalues 1, 2, and 6. It is easily verified that $(1/\sqrt{12})m_\varphi$ is not a Kreĭn contraction but $(1/\sqrt{20})m_\varphi$ (among other multiples) is a Kreĭn contraction.

As in the case of Theorem 2, there is an easy analogue in more general Kreĭn spaces with an identical proof.

We are able to get more precise information if we have other boundary information. In particular, if one boundary point is mapped to another boundary point, then the multiplier is uniquely determined, but this multiplier is not as easily determined.

Theorem 5. *Let φ be a linear fractional self-map of $\mathbb{B}_N$ with no boundary fixed point. Let d be a vector associated to the interior fixed point. Assume also that m_φ has been chosen so that the eigenvalue λ_1 associated with d is positive. If there is a vector $x \in \mathbb{C}^{N+1}$ such that $[x,x] = [m_\varphi x, m_\varphi x] = 0$, then tm_φ is a Kreĭn contraction where*

$$t = \sqrt{\frac{\operatorname{Re}([d,x])}{\lambda_1 \operatorname{Re}([d, m_\varphi x])}}.$$

Proof. This proof will follow in a manner similar to that in the boundary fixed point case. Since φ is assumed to be a self-map of the ball, there exists $t > 0$ such that tm_φ is a Kreĭn contraction. For simplicity of notation, we will again define $y = m_\varphi x$. We let $v = sd + (1-s)x$ so that $m_\varphi v = s\lambda_1 d + (1-s)y$ and

$$[m_\varphi v, m_\varphi v] = s^2\lambda_1^2[d,d] + 2s(1-s)\lambda_1 \operatorname{Re}([d,y])$$

since, by our hypotheses, $[y,y] = 0$. Similarly,

$$[v,v] = s^2[d,d] + 2s(1-s)\operatorname{Re}([d,x])$$

since $[x,x] = 0$. By neglecting the quadratic terms, we have

$$t^2 s(1-s)\lambda_1 \operatorname{Re}[d,y] \le s(1-s)\operatorname{Re}[d,x]$$

for all sufficiently small s. By taking small positive and negative values for s, we obtain the desired expression for t. □

In practice, finding boundary values which map to other boundary values is not an easy process, so unless such an occurrence is known, one will usually need another means to identify this case. Recall that the statement "tm_φ is a Kreĭn contraction" is equivalent to the statement "$J - t^2 m_\varphi^* J m_\varphi$ has no negative eigenvalues". This latter matrix is self-adjoint, so all eigenvalues are real. We can make use of this equivalence to find this type of boundary behavior and find the appropriate multiple to test. If $\chi(\lambda, t)$ is the characteristic polynomial of $J - t^2 m_\varphi^* J m_\varphi$, then the coefficient of λ^n is a polynomial in t^2. If the multiplier for which tm_φ is a Kreĭn contraction is unique, then one of these polynomials in t^2, generally the coefficient of λ^0, must have a double root. To see why this is true, we consider the case $N = 2$. If there are no negative eigenvalues, the coefficient of λ^0 must be non-positive, the coefficient of λ^1 must be non-negative, and the coefficient of λ^2 must be non-positive. If there is no double root of any of these, then there will be an open interval on which all are satisfied or there will be no value of t^2 for which all are satisfied.

Example 6. If φ is a linear fractional map with associated matrix of the form

$$m_\varphi = \begin{pmatrix} 0 & -\frac{k}{2} & \frac{k}{2} \\ b & 0 & 0 \\ 0 & 0 & k \end{pmatrix},$$

then $\varphi((0,-1)) = (1,0)$ and φ has no boundary fixed points. Its eigenvalues are $\{k, \pm\sqrt{-bk/2}\}$. The coefficient of λ^0 in the characteristic polynomial of $J - t^2 m_\varphi^* J m_\varphi$ is $(1-t^2b^2)(1-t^2k^2a)^2$ so $t^2 = 1/(k^2a)$ is the right multiple to check.

Note that the presence of a double root is also true in the case of boundary fixed points, but finding the exact solution in this manner is more likely to be difficult than using the method of Theorem 1. This methodology also suggests a method

to guess a multiple when there is no boundary fixed point. One can look for an interval in which the various coefficients of the characteristic polynomial have the correct sign.

References

[1] D. Alpay, A. Dijksma, J. Rovnyak, and H. de Snoo, *Schur functions, operator colligations, and reproducing kernel Pontryagin spaces*, Birkhäuser Verlag, Basel, 1997. MR **2000a:**47024

[2] C. Bisi and F. Bracci, *Linear fractional maps of the unit ball: A geometric study*, preprint, 2000.

[3] J. Bognár, *Indefinite inner product spaces*, Springer-Verlag, New York, 1974, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 78. MR **57:**7125

[4] F. Bracci, *On the geometry at the boundary of holomorphic self-maps of the unit ball of* $\mathbf{c}^n$, Complex Variables Theory Appl. **38** (1999), no. 3, 221–241. MR **2000b:**32038

[5] C. C. Cowen and B. D. MacCluer, *Composition Operators on Spaces of Analytic Functions*, CRC Press, Boca Raton, 1995. MR **97i:**47056

[6] ———, *Schroeder's equation in several variables*, preprint, August 1999.

[7] ———, *Linear fractional maps of the ball and their composition operators*, Acta Sci. Math. (Szeged) **66** (2000), no. 1-2, 351–376. MR **2001g:**47041

[8] D. Crosby, *A breakdown of linear fractional maps of the ball*, unpublished notes from research as an undergraduate, 1996.

[9] M. A. Dritschel and J. Rovnyak, *Extension theorems for contraction operators on Kreĭn spaces*, Extension and interpolation of linear operators and matrix functions, Birkhäuser, Basel, 1990, pp. 221–305. MR **92m:**47068

[10] J. W. Helton, J. A. Ball, C. R. Johnson, and J. N. Palmer, *Operator theory, analytic functions, matrices, and electrical engineering*, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1987. MR **89f:**47001

[11] M. G. Kreĭn and Ju. L. Šmul′jan, *Plus-operators in a space with indefinite metric*, Amer. Math. Soc. Transl. (2) **85** (1969), 93–113.

[12] V. P. Potapov, *Linear fractional transformations of matrices*, Amer. Math. Soc. Transl. (2) **138** (1988), 21–35.

[13] B. Schwarz and A. Zaks, *Non-Euclidean motions in projective matrix spaces*, Linear Algebra Appl. **137/138** (1990), 351–361. MR **92a:**51025

[14] J. H. Shapiro, *Composition Operators and Classical Function Theory*, Springer-Verlag, New York, 1993. MR **94k:**47049

[15] Ju. L. Šmul′jan, *General linear-fraction maps of operator balls*, Siberian Math. J. **19** (1978), no. 2, 293–298. MR **58:**12463

Department of Mathematics, Purdue University, West Lafayette, Indiana 47907

E-mail address: `richman@math.purdue.edu`

Current address: *as of August 11, 2002*: Department of Mathematics, Bucknell University, Lewisburg, Pennsylvania 17837

E-mail address: `arichman@bucknell.edu`