NORMS OF COMPOSITION OPERATORS WITH RATIONAL SYMBOL

SEAN EFFINGER-DEAN, ALAN JOHNSON, JOSEPH REED, AND JONATHAN SHAPIRO

Abstract. We compute the norms of composition operators with rational symbols that satisfy certain properties, extending Christopher Hammond's methods on operators with linear fractional symbols. This leads to a host of new examples of composition operators whose norms are calculable.

1. Introduction

Let $\mathbb{D}$ be the open unit disk in the complex plane. The Hardy space H^2 is the familiar Hilbert space of analytic functions on $\mathbb{D}$ with square-summable Taylor coefficients.

For φ an analytic self-map of $\mathbb{D}$, C_φ denotes the *composition operator* defined by $C_\varphi f = f \circ \varphi$. Littlewood's Subordination Principle, which can be found in [7], guarantees that C_φ is a bounded operator on H^2. We are interested in calculating the norm of C_φ. This is a difficult problem in general, so we restrict our attention to the case when φ is rational. We now introduce several concepts that we will use frequently in this paper.

Definition 1.1. For $z \in \mathbb{D}$, let $K_z : \mathbb{D} \to \mathbb{C}$ be given by

$$K_z(\zeta) = \frac{1}{1 - \overline{z}\,\zeta}.$$

It is easy to check that $K_z \in H^2$ and that this function has the property that for any $f \in H^2$, $\langle f, K_z \rangle = f(z)$. For this reason K_z is called the *reproducing kernel* at z.

Also useful in the study of analytic functions on the disk is the following:

Definition 1.2. An analytic $\varphi : \mathbb{D} \to \mathbb{D}$ is called *inner* if $|\varphi(e^{i\theta})| = 1$ for almost every $\theta \in [0, 2\pi]$.

We now define a simple and fundamental class of inner functions.

Definition 1.3. For $z \in \mathbb{D}$, the function $\Phi_z : \mathbb{D} \to \mathbb{D}$ is defined as

$$\Phi_z(\zeta) = \frac{\zeta - z}{1 - \overline{z}\,\zeta}.$$

Note that Φ_z is an automorphism of the disk that vanishes at z.

Date: December 14, 2005.

1991 *Mathematics Subject Classification.* 47B33.

Key words and phrases. Composition Operator, Hardy Space, Norm.

This material is based upon work partially supported by the National Science Foundation under Grant No. DMS-0353622 as part of the REU program at Cal Poly, San Luis Obispo.

Definition 1.4. An *isometry* is an operator A on a Hilbert space $\mathcal{H}$ with the property that for all $f, g \in \mathcal{H}$, $\langle Af, Ag\rangle = \langle f, g\rangle$. If C_φ is an isometry, we say that φ is an *isometry-inducing function.*

Our goal is to calculate the exact norm of composition operators whose symbols are in a certain special class of rational functions. At present, there is a very limited collection of self-maps φ for which $\|C_\varphi\|$ is known exactly. These include inner functions, for which $\|C_\varphi\| = \sqrt{\frac{1+|\varphi(0)|}{1-|\varphi(0)|}}$, constant maps $\varphi \equiv a$, for which $\|C_\varphi\| = \sqrt{\frac{1}{1-|a|^2}}$, and even all linear maps $\varphi(z) = sz + t$ with $|t| < 1$ and $|s| + |t| \leq 1$: In this case (see [2] or [3, p. 324]),

$$\|C_\varphi\| = \sqrt{\frac{2}{1 + |s|^2 - |t|^2 + \sqrt{(1 - |s|^2 + |t|^2)^2 - 4|t|^2}}}.$$

C. Hammond, in [4] and [5], and, with P. Bourdon, E. Fry, and C. Spofford in [1], developed techniques to compute the norm of a composition operator, in many cases, with linear fractional symbol. In this paper we extend the methods of these earlier papers to allow us to compute composition operator norms when the symbol is in a special class of (higher order) rational functions.

If $\varphi = \tau \circ \psi$ are all analytic self-maps of the disk, then $C_\varphi = C_\psi C_\tau$. If ψ is an isometry-inducing function then it is clear that $\|C_\varphi\| = \|C_\tau\|$. The set of isometry-inducing functions is precisely the set of inner functions which fix the origin, see [6] or [3, pp. 123-124]. This allows us to extend our collection of composition operators with calculable norms in a somewhat trivial way, for example: Let $\varphi(z) = \frac{z^2+1}{2}$. We can write $\varphi = \tau \circ \psi$ for $\tau(z) = \frac{z+1}{2}$ and $\psi(z) = z^2$, an isometry-inducing function. We then compute $\|C_\varphi\| = \|C_\tau\| = \sqrt{2}$ by the formula above.

When we find new examples of φ with calculable norm, we will prove that there do not exist simpler τ and isometry-inducing ψ with $\varphi = \tau \circ \psi$.

For notational convenience, we introduce the following function:

Definition 1.5. $\rho : \mathbb{C}^* \to \mathbb{C}^*$ (where $\mathbb{C}^*$ denotes the extended complex plane) is defined by $\rho(z) = 1/\overline{z}$. Note that $\rho^{-1} = \rho$ and for $z \in \partial\mathbb{D}$, $\rho(z) = z$.

2. Rational Functions with Calculable Composition Operator Norms

The main reason we restrict ourselves to rational φ is that C_φ^* can then be written in terms of an integral of a meromorphic function. This allows us to investigate the behavior of $C_\varphi^* C_\varphi$ more closely and, in some cases, to compute its eigenvalues. As long as C_φ is norm-attaining, $\|C_\varphi^* C_\varphi\| = \|C_\varphi\|^2$ is an eigenvalue of $C_\varphi^* C_\varphi$. We will require the following lemmas before we prove the main result. These lemmas and the ensuing proofs appear in Hammond's papers, [4] and [5], but we would like to include them here for completeness.

Lemma 2.1. *Let T be a self-adjoint operator on a Hilbert space $\mathcal{H}$ with a closed subspace W that is invariant under T. Then for any eigenvalue of T, there exists a corresponding eigenfunction in W or in $W^\perp$.*

Proof. Let λ be an eigenvalue of T with corresponding eigenfunction g. Then there is a unique decomposition $g = g_1 + g_2$, with $g_1 \in W$ and $g_2 \in W^\perp$. Then

$$Tg = \lambda\, g = \lambda\, g_1 + \lambda\, g_2.$$

Also, $Tg = Tg_1 + Tg_2$. The subspace $W^\perp$ is also invariant under T because the operator is self-adjoint. Hence $Tg_1 \in W$ and $Tg_2 \in W^\perp$. Since the decomposition of Tg is unique, we have $Tg_1 = \lambda\, g_1$ and $Tg_2 = \lambda\, g_2$. Because either g_1 or g_2 is non-zero, at least one represents an eigenfunction of T with eigenvalue λ. □

Lemma 2.2. *Let T be a bounded operator on a Hilbert space $\mathcal{H}$. Let g be a maximizing vector for T^*T, i.e., a function with the property that $\|T^*Tg\| = \|T^*T\|\,\|g\|$. Then g is a maximizing vector for T.*

Proof. We have the well-known identities $\|T^*\| = \|T\|$ and $\|T^*T\| = \|T\|^2$. Therefore we have

$$\|T\|^2\|g\| = \|T^*Tg\| \le \|T^*\|\,\|Tg\| = \|T\|\,\|Tg\|.$$

Hence $\|Tg\| \ge \|T\|\,\|g\|$. Clearly, $\|Tg\| \le \|T\|\,\|g\|$, so $\|Tg\| = \|T\|\,\|g\|$. Therefore, g is a maximizing vector for T. □

Lemma 2.3. *Let $\varphi : \mathbb{D} \to \mathbb{D}$ be a non-inner analytic function and let g be a maximizing vector for C_φ. Then g is non-vanishing on $\mathbb{D}$.*

Proof. Suppose that g vanishes at the point $z_0 \in \mathbb{D}$. Then let $h = g/B_{z_0}$, where B_{z_0} is the Blaschke factor which vanishes at z_0. Then h is analytic, and for $z \in \partial\mathbb{D}$, $|h(z)| = |g(z)|$, so $\|h\| = \|g\|$. Also, since we may assume that g is not identically zero, we have $|h(z)| > |g(z)|$ almost everywhere in $\mathbb{D}$. Because φ is non-inner, $|h(\varphi(z))| > |g(\varphi(z))|$ on a subset of $\partial\mathbb{D}$ which has positive measure. Hence $\|C_\varphi h\| > \|C_\varphi g\|$, contradicting the assumption that g is norm-attaining. □

Theorem 2.4. *Suppose $\varphi : \mathbb{D} \to \mathbb{D}$ extends to a non-inner rational function on $\mathbb{C}^*$ and assume that C_φ is norm-attaining. Let $A = \{\zeta_k\}_{k=1}^n \subset \mathbb{D}$ denote the set of roots of the function $h(\zeta) = \zeta\left(1 - \varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right)$. Suppose that each of these roots has multiplicity 1 and that $\varphi(A) \subset \{0, \varphi(0)\}$. Now let*

$$a_1 = \sum_{\varphi(\zeta_k)=0} \operatorname*{Res}_{\zeta=\zeta_k} \frac{1}{\zeta\left(1 - \varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right)}$$

$$a_2 = \sum_{\varphi(\zeta_k)=\varphi(0)} \operatorname*{Res}_{\zeta=\zeta_k} \frac{1}{\zeta\left(1 - \varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right)}\,.$$

Then $\lambda = \|C_\varphi\|^2$ is the greatest solution to the following quadratic equation:

$$\lambda^2 - a_2\lambda - a_1 = 0.$$

Proof. In order to compute C_φ^*, we use the kernel functions of the Hardy space. Note that, for any $f \in H^2$, $(C_\varphi^* f)(z) = \langle C_\varphi^* f, K_z\rangle = \langle f, C_\varphi K_z\rangle$. Hence we have the following expression for $C_\varphi^* C_\varphi$:

$$(C_\varphi^* C_\varphi f)(z) = \langle C_\varphi f, C_\varphi K_z\rangle = \frac{1}{2\pi}\int_0^{2\pi} \frac{f(\varphi(e^{i\theta}))}{1 - z\,\overline{\varphi(e^{i\theta})}}\,d\theta.$$

We now change variables, letting $\zeta = e^{i\theta}$,

$$(C_\varphi^* C_\varphi f)(z) = \frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{f(\varphi(\zeta))}{\zeta(1 - z\,\overline{\varphi(\zeta)})}\,d\zeta.$$

Recall that, for $\zeta \in \partial\mathbb{D}$, $\zeta = \rho(\zeta)$, so the expression can be rewritten as

$$(C_\varphi^* C_\varphi f)(z) = \frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{f(\varphi(\zeta))}{\zeta\left(1 - z\,\overline{(\varphi\circ\rho)(\zeta)}\right)}\, d\zeta.$$

Since φ is rational, $\overline{\varphi\circ\rho}$ is also rational on $\partial\mathbb{D}$. Hence the integrand can be written as a meromorphic function on $\mathbb{D}$. Therefore the integral can be computed using residues. This gives us the following:

$$\begin{aligned}(C_\varphi^* C_\varphi\, f)(\varphi(0)) &= \frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{f(\varphi(\zeta))}{\zeta\left(1 - \varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right)}\, d\zeta \\ &= \sum_{k=1}^{n} \operatorname*{Res}_{\zeta=\zeta_k} \frac{f(\varphi(\zeta))}{\zeta\left(1 - \varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right)} \\ &= \sum_{k=1}^{n} f(\varphi(\zeta_k)) \operatorname*{Res}_{\zeta=\zeta_k} \frac{1}{\zeta\left(1 - \varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right)}\ ,\end{aligned}$$

where we may take $f(\varphi(\zeta_k))$ outside the residue expression because each ζ_k is a pole of multiplicity 1. Using the values of a_1 and a_2 stated in the proposition, we have

$$(C_\varphi^* C_\varphi\, f)(\varphi(0)) = a_1\, f(0) + a_2\, f(\varphi(0)). \tag{1}$$

We now use these identities to demonstrate exactly how $C_\varphi^* C_\varphi$ acts on the kernel functions K_0 and $K_{\varphi(0)}$. Since $K_0(z) = 1$ is a constant function, it is unchanged by C_φ, so

$$C_\varphi^* C_\varphi K_0 = C_\varphi^* K_0 = K_{\varphi(0)}.$$

For any $z \in \mathbb{D}$,

$$\begin{aligned}(C_\varphi^* C_\varphi K_{\varphi(0)})(z) &= \langle C_\varphi^* C_\varphi K_{\varphi(0)}, K_z\rangle = \overline{\langle C_\varphi^* C_\varphi K_z, K_{\varphi(0)}\rangle} \\ &= \overline{(C_\varphi^* C_\varphi K_z)(\varphi(0))} = \overline{a_1 K_z(0) + a_2 K_z(\varphi(0))} \\ &= \overline{a_1}\, K_0(z) + \overline{a_2}\, K_{\varphi(0)}(z),\end{aligned}$$

where the last line uses equation (1). Let $W = \operatorname{Span}\left\{K_0, K_{\varphi(0)}\right\}$. Then the above identities show that W is invariant under $C_\varphi^* C_\varphi$. Let g be a maximizing eigenvector for $C_\varphi^* C_\varphi$, i.e., an eigenvector whose eigenvalue is the norm. By Lemma 2.2, g is also a maximizing vector for C_φ. Further, by Lemma 2.1, we may assume that $g \in W$ or that $g \in W^\perp$. If $g \in W^\perp$, then it vanishes at 0 and $\varphi(0)$, contradicting Lemma 2.3. Hence $g \in W$, so $g = c_1 K_0 + c_2 K_{\varphi(0)}$ for some $c_1, c_2 \in \mathbb{C}$, not both zero. Because of our identities for $C_\varphi^* C_\varphi K_0$ and $C_\varphi^* C_\varphi K_{\varphi(0)}$, and since g is an eigenfunction, c_1 and c_2 must satisfy

$$\lambda \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \overline{a_1} & \overline{a_2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Therefore, the set of eigenvalues of $C_\varphi^* C_\varphi$ on W is precisely the set of solutions to

$$\begin{vmatrix} -\lambda & 1 \\ \overline{a_1} & \overline{a_2} - \lambda \end{vmatrix} = 0.$$

By taking the conjugate of both sides and noting that $\lambda \in \mathbb{R}$, this is equivalent to the equation $\lambda^2 - a_2\lambda - a_1 = 0$. Hence the greatest solution to this equation is the greatest eigenvalue of $C_\varphi^* C_\varphi$, and therefore is $\|C_\varphi^* C_\varphi\| = \|C_\varphi\|^2$. □

Example 2.5. We consider an example of a symbol φ which satisfies the conditions of Theorem 2.4. Let

$$\varphi(z) = \frac{64 + 60\,z - 136\,z^2}{256 + 15\,z - 94\,z^2}\ .$$

It is easy to check that this is an analytic self-map of $\mathbb{D}$ with $\|\varphi\|_\infty < 1$. Therefore C_φ is compact and hence norm-attaining. We then have

$$h(\zeta) = \zeta\left(1 - \varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right) = \frac{60\zeta - 240\zeta^3}{94 - 15\zeta - 256\zeta^2} = \frac{60\zeta(1-2\zeta)(1+2\zeta)}{94 - 15\zeta - 256\zeta^2}\ ,$$

so the set of roots is $A = \{0, -\frac{1}{2}, \frac{1}{2}\}$. Each of these roots has multiplicity 1, as desired, and $\varphi(A) = \{0, \frac{1}{4}\} = \{0, \varphi(0)\}$. Then $a_1 = -\frac{5}{16}$ and $a_2 = \frac{331}{240}$. By taking the largest root of the quadratic equation obtained from a_1 and a_2, we see that

$$\|C_\varphi\|^2 = \frac{331 + \sqrt{37561}}{480} \approx 1.09335.$$

3. Comparison with Hammond's Theorem

C. Hammond's theorem, from [5, Theorem 5.5], tells us

Theorem 3.1 (Hammond)**.** *Let $\varphi : \mathbb{D} \to \mathbb{D}$ be a linear fractional map, with $\varphi(z) \neq az$. Suppose that $\tau_n(\varphi(0)) = 0$ for some integer $n \geq 0$; then $\|C_\varphi\|^2$ is the largest zero of the polynomial*

$$p(\lambda) = \lambda^{n+1} - \sum_{k=0}^{n} \chi\left(\tau_k\left(\varphi(0)\right)\right)\left[\prod_{m=0}^{k-1} \psi\left(\tau_m\left(\varphi(0)\right)\right)\right]\lambda^{n-k},$$

and the elements on which C_φ attains its norm are linear combinations of the kernel functions $\left\{K_{\tau_j(\varphi(0))}\right\}_{j=0}^{n}$.

Here, we use, for the linear fractional map $\varphi(z) = \frac{az+b}{cz+d}$, the auxiliary functions $\sigma(z) = \frac{\overline{a}z-\overline{c}}{-\overline{b}z+\overline{d}}$, $\tau(z) = \varphi\left(\sigma(z)\right)$, and

$$\psi(z) = \frac{\left(\overline{ad} - \overline{bc}\right)z}{\left(\overline{a}z - \overline{c}\right)\left(-\overline{b}z + \overline{d}\right)} \quad \text{and} \quad \chi(z) = \frac{\overline{c}}{-\overline{a}z + \overline{c}}.$$

In the special case when $n = 1$, Hammond's theorem tells us that if $\varphi : \mathbb{D} \to \mathbb{D}$ is a linear fractional map with $\tau\left(\varphi(0)\right) = 0$, then $\|C_\varphi\|^2$ is the largest zero of the polynomial

$$p(\lambda) = \lambda^2 - \chi\left(\varphi(0)\right)\lambda - \psi\left(\varphi(0)\right). \tag{2}$$

For $\varphi(z) = \frac{az+b}{cz+d}$, the condition $\tau\left(\varphi(0)\right) = 0$ is equivalent to $\frac{\overline{a}b-\overline{c}d}{b\overline{b}-d\overline{d}} = \frac{b}{a}$, and we can compute the coefficients in the quadratic polynomial above: $\chi\left(\varphi(0)\right) = \frac{\overline{c}d}{\overline{c}d-\overline{a}b}$ and $\psi\left(\varphi(0)\right) = \frac{\left(\overline{a}\overline{d}-\overline{b}\overline{c}\right)bd}{\left(\overline{a}b-\overline{c}d\right)\left(d\overline{d}-b\overline{b}\right)}$.

To compare the above computation of the composition operator norm with that using Theorem 2.4, we first must note that for the above function φ, $\varphi\left(0\right) = \frac{b}{d}$, and the roots of $h(\zeta) = \zeta\left(1 - \varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right) = \zeta\left(\frac{\left(d\overline{c}-b\overline{a}\right)+\left(d\overline{d}-b\overline{b}\right)\zeta}{d\overline{c}+d\overline{d}\zeta}\right)$ are the elements of the set $A = \left\{0, \frac{\overline{a}b-\overline{c}d}{d\overline{d}-b\overline{b}}\right\} = \left\{0, -\frac{b}{a}\right\}$. Since $\varphi\left(-\frac{b}{a}\right) = 0$, it is then easy to see that

$\varphi(A) = \varphi\left\{0, -\frac{b}{a}\right\} = \{\varphi(0), 0\}$, so the hypotheses of Theorem 2.4 hold. Theorem 2.4 then tells us that $\|C_\varphi\|^2$ is the largest zero of the polynomial

$$p(\lambda) = \lambda^2 - a_2\lambda - a_1. \tag{3}$$

We can compute

$$\begin{aligned} a_2 &= \sum_{\varphi(\zeta_k)=\varphi(0)} \operatorname*{Res}_{\zeta=\zeta_k} \frac{1}{\zeta\left(1-\varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right)} \\ &= \operatorname*{Res}_{\zeta=0} \frac{1}{\zeta\left(\frac{(d\bar{c}-b\bar{a})+\left(d\bar{d}-b\bar{b}\right)\zeta}{d\bar{c}+d\bar{d}\zeta}\right)} = \frac{\bar{c}d}{\bar{c}d-\bar{a}b} \end{aligned}$$

and

$$\begin{aligned} a_1 &= \sum_{\varphi(\zeta_k)=0} \operatorname*{Res}_{\zeta=\zeta_k} \frac{1}{\zeta\left(1-\varphi(0)\,\overline{(\varphi\circ\rho)(\zeta)}\right)} \\ &= \operatorname*{Res}_{\zeta=-\frac{b}{a}} \frac{1}{\zeta\left(\frac{(d\bar{c}-b\bar{a})+\left(d\bar{d}-b\bar{b}\right)\zeta}{d\bar{c}+d\bar{d}\zeta}\right)} \\ &= \frac{\left(\bar{a}\bar{d}-\bar{b}\bar{c}\right)bd}{(\bar{a}b-\bar{c}d)\left(d\bar{d}-b\bar{b}\right)} \quad \text{(after some messy algebra).} \end{aligned}$$

This tells us that the coefficients in the two polynomials from equations (2) and (3) are identical, and thus the computations of the composition operator norms are the same as well. The result is that the "$n = 1$" version of Hammond's theorem is a special case of our Theorem 2.4.

4. When can φ be written as a composition of simpler self-maps?

It is worth noting that the φ in the example above cannot be expressed as a linear fractional map composed with an isometry-inducing function. This is a consequence of the following proposition, which characterizes precisely when φ can be expressed as such.

Proposition 4.1. *Suppose $\varphi : \mathbb{D} \to \mathbb{D}$ extends to a rational function on $\mathbb{C}^*$, and fix $c_1 \in \mathbb{D}$. Let R be the set of roots of $(\rho\circ\varphi\circ\rho)(z) - c_1$ and suppose that each of these roots has multiplicity* 1*. Then the following two conditions are equivalent:*

1. *There exists $c_2 \in \mathbb{D}$ such that for all $z \in R$, $\varphi(z) = c_2$.*
2. *$\varphi = \ell \circ \psi$ for some linear fractional $\ell : \mathbb{D} \to \mathbb{D}$ and inner ψ with $\psi(0) = 0$.*

Proof. We first show that condition 1 implies condition 2. Suppose φ has degree d. Then R has precisely d elements since $(\rho\circ\varphi\circ\rho)(z) - c_1$ is also degree d. Note that $\varphi(z) - \rho(c_1) = 0$ whenever $z \in \rho(R)$. Because $\rho(R)$ has d distinct elements, $\rho(R)$ is precisely the set of roots of $\varphi(z) - \rho(c_1)$. By similar reasoning, the set of roots of $\varphi(z) - c_2$ is precisely R (based on condition 1). Also note that for all $z \in R$, $\rho(z) \notin \overline{\mathbb{D}}$ because $\varphi(\rho(z)) = \rho(c_1) \notin \overline{\mathbb{D}}$. Hence $z \in \mathbb{D}$. We define

$$g(z) = \frac{z - c_2}{z - \rho(c_1)}\ .$$

We also define

$$\Psi = \prod_{z \in A} \Phi_z.$$

The set of roots of $g \circ \varphi$ is precisely R and the set of poles is precisely $\rho(R)$. Note that these coincide exactly with the roots and poles of Ψ. Since both $g \circ \varphi$ and Ψ are rational functions with identical zeros and poles, Ψ is a scalar multiple of $g \circ \varphi$; say $g \circ \varphi = \kappa \Psi$, with $\kappa \in \mathbb{C} - \{0\}$. Note that g is non-constant (since $g(\rho(c_1)) = \infty$ and $g(c_2) = 0$), and hence has a well-defined linear-fractional inverse g^{-1}. Let $\ell = g^{-1} \circ \kappa \Phi_{\Psi(0)}^{-1}$ and let $\psi = \Phi_{\Psi(0)} \circ \Psi$. Then $\varphi = \ell \circ \psi$. Note that ψ is an inner function and that $\psi(0) = \Phi_{\Psi(0)}(\Psi(0)) = 0$, as desired. Also, ℓ is a linear fractional map since it is the composition of linear fractional maps. This function ℓ must be a self-map of the disk since (using the fact that ψ is surjective) $\ell(\mathbb{D}) = (\ell \circ \psi)(\mathbb{D}) = \varphi(\mathbb{D}) \subset \mathbb{D}$.

We now prove that condition 2 implies condition 1. First suppose that ℓ is non-constant. Then ℓ^{-1} is well-defined and $\psi = \ell^{-1} \circ \varphi$, so ψ is rational. Because ψ is inner and rational, $|\psi(z)| = 1$ for all $z \in \partial\mathbb{D}$. Hence $\psi(z) = (\rho \circ \psi \circ \rho)(z)$ for all $z \in \partial\mathbb{D}$. Since ψ and $\rho \circ \psi \circ \rho$ are rational functions which agree on $\partial\mathbb{D}$, they agree everywhere. So $(\rho \circ \varphi \circ \rho)(z) = c_1$ if and only if $(\psi \circ \rho)(z) = (\ell^{-1} \circ \rho)(c_1)$. This is true if and only if $(\rho \circ \psi \circ \rho)(z) = \psi(z) = (\rho \circ \ell^{-1} \circ \rho)(c_1)$. Letting $c_2 = (\ell \circ \rho \circ \ell^{-1} \circ \rho)(c_1)$, this equation becomes $\varphi(z) = c_2$. Finally, we know that $c_2 \in \mathbb{D}$ because $A \subset \mathbb{D}$ and $\varphi : \mathbb{D} \to \mathbb{D}$.

Now suppose that ℓ is constant. Then φ is constant, so say $\varphi = \kappa$, with $\kappa \in \mathbb{D}$. Then for all $z \in \mathbb{C}^*$, $(\rho \circ \varphi \circ \rho)(z) = \rho(\kappa)$. Since $c_1 \in \mathbb{D}$ and $\rho(\kappa) \notin \mathbb{D}$, $R = \emptyset$. Therefore for any $c_2 \in \mathbb{D}$, condition 1 is vacuously true. □

Corollary 4.2. *Suppose φ satisfies the hypotheses of Theorem 2.4. We may apply Proposition 4.1 (letting $c_1 = \varphi(0)$) to show that $\varphi = \ell \circ \psi$ for some linear fractional $\ell : \mathbb{D} \to \mathbb{D}$ and isometry-inducing ψ if and only if φ maps each nonzero element in the set A to 0 (in which case $c_2 = 0$ above).*

Proof. In order to satisfy Proposition 4.1, φ must map all of the nonzero elements of A, i.e., roots of $1 - \varphi(0)\overline{(\varphi \circ \rho)(\zeta)}$, to 0 or all to $\varphi(0)$. Assuming $\varphi(0) \neq 0$, we prove that the second case is impossible by contradiction. If φ has degree d, then $1 - \varphi(0)\overline{(\varphi \circ \rho)(\zeta)}$ has d roots. By the hypotheses of Theorem 2.4, these d roots are distinct. Because φ sends each of these roots to $\varphi(0)$, 0 is one of the roots (since $\varphi(\zeta) = \varphi(0)$ has at most d distinct solutions, one of which is $\zeta = 0$). This contradicts the hypotheses of Theorem 2.4 because then 0 is a root of $\zeta(1 - \varphi(0)\overline{(\varphi \circ \rho)(\zeta)})$ with multiplicity 2. Hence φ equals a linear fractional map composed with an isometry-inducing function if and only if $\varphi(A - \{0\}) = \{0\}$. □

In Example 2.5, $\frac{1}{2} \in A$, and $\varphi(\frac{1}{2}) = \frac{1}{4} = \varphi(0)$, confirming that this φ cannot be expressed as a linear fractional map composed with an isometry-inducing function.

5. Generating Examples

One may easily construct a variety of other non-trivial examples for Theorem 2.4. We show how to construct an example of degree d. Fix a set $\{\zeta_k\}_{k=1}^{d} \subset \mathbb{D} - \{0\}$, with $\zeta_j \neq \zeta_k$ for $j \neq k$, and fix $\varphi(0)$. Also designate which ζ_k's are mapped to 0 by φ and which are mapped to $\varphi(0)$. Let

$$\varphi(z) = \frac{\varphi(0) + \sum_{k=1}^{d} a_k z^k}{1 + \sum_{k=1}^{d} b_k z^k} .$$

Note that the equation $1-\varphi(0)\overline{(\varphi \circ \rho)(\zeta_k)} = 0$ can be rewritten as a linear equation in the a_k's and b_k's. The same is true for the equations $\varphi(\zeta_k) = 0$ and $\varphi(\zeta_k) = \varphi(0)$ (for each k, one of these two equations holds). Hence we have $2d$ linear equations and $2d$ unknowns, so we may solve for the coefficients $\{a_k, b_k\}_{k=1}^{d}$, thereby deriving an expression for φ. The only remaining concern is whether φ is a self-map of the disk. As it turns out, placing the ζ_k's close enough to the boundary $\partial\mathbb{D}$ and $\varphi(0)$ close enough to 0 solves this problem.

Example 5.1. We consider an example of the above process when $d = 3$. Let $\{\zeta_k\}_{k=1}^{3} = \left\{\frac{1}{2}, \frac{2}{3}, -\frac{2}{3}\right\}$ (so $A = \left\{0, \frac{1}{2}, \frac{2}{3}, -\frac{2}{3}\right\}$) and let φ send all of these ς's to 0. Also let $\varphi(0) = \frac{1}{6}$. Then we have

$$\varphi(z) = \frac{216 - 432\, z - 486\, z^2 + 972\, z^3}{1296 - 702\, z - 641\, z^2 + 442\, z^3} .$$

It is easy to check that φ is a self-map of $\mathbb{D}$. Corollary 4.2 guarantees that this is a linear fractional map composed with an isometry-inducing function, and indeed, if we let

$$\ell(z) = \frac{108 - 486\, z}{648 - 221\, z} \qquad \text{and} \qquad \psi(z) = z\,\frac{54 + 65\, z - 154\, z^2}{154 - 65\, z - 54\, z^2} ,$$

then $\varphi = \ell \circ \psi$. Both φ and ℓ satisfy the conditions of Theorem 2.4. This means that we can use the methods of Theorem 2.4 directly on φ, or, alternatively, use the methods of Theorem 2.4 on ℓ. Doing either with a simple calculation (in both cases the $a_1 = -\frac{11}{20}$ and $a_2 = \frac{221}{140}$), we see that $\|C_\varphi\|^2 = \|C_\ell\|^2 = \frac{1}{280}(221 + \sqrt{5721}) \approx 1.05942$.

Example 5.2. If we assign the same values to $\{\zeta_k\}_{k=1}^{3} = \{\frac{1}{2}, \frac{2}{3}, -\frac{2}{3}\}$, but let $\varphi(0) = \frac{1}{20}$ and choose a φ which sends $\frac{1}{2}$ to $\varphi(0)$, and still have $\varphi(\frac{2}{3}) = 0$ and $\varphi(-\frac{2}{3}) = 0$, then we come up with

$$\varphi(z) = \frac{336 + 352\, z - 756\, z^2 - 792\, z^3}{6720 - 3334\, z - 3017\, z^2 + 1450\, z^3} .$$

It is easy to check that this φ is also a self-map of $\mathbb{D}$. Corollary 4.2 shows us that unlike our previous example, this map φ cannot be expressed as a linear fractional map composed with an isometry-inducing function. Using Theorem 2.4, we see that $\|C_\varphi\|^2 = \frac{1}{156408}(82365 + \sqrt{5543677785}) \approx 1.00264$.

6. A More General Result

Theorem 6.1. *Suppose $\varphi : \mathbb{D} \to \mathbb{D}$ extends to a non-inner rational function on $\mathbb{C}^*$ and assume that C_φ is norm-attaining. Say there exist nonempty sets $A = \{\zeta_i\}_{i=1}^{m} \subset \mathbb{D}$ and $B = \{z_j\}_{j=1}^{n} \subset \mathbb{D}$ with the following properties:*

1. Each root of $\zeta\left(1 - z_k\,\overline{(\varphi \circ \rho)(\zeta)}\right)$ has multiplicity 1 and is an element of A.

2. $\varphi(A) \subset B$.

Let M be the $n \times n$ matrix with entries

$$m_{jk} = \sum_{\varphi(\zeta_i)=z_j} \operatorname*{Res}_{\zeta=\zeta_i} \frac{1}{\zeta\left(1 - z_k\,\overline{(\varphi\circ\rho)(\zeta)}\right)} .$$

Then $\|C_\varphi\|^2$ is the greatest eigenvalue of M.

Proof. We follow essentially the same argument as the proof to Theorem 2.4. For $1 \le k \le n$ and for any $f \in H^2$, using conditions 1 and 2 from the statement of the proposition, we have

$$\begin{aligned}(C_\varphi^* C_\varphi f)(z_k) &= \frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{f(\varphi(\zeta))}{\zeta\left(1 - z_k\,\overline{(\varphi\circ\rho)(\zeta)}\right)} = \sum_{i=1}^{m} \operatorname*{Res}_{\zeta=\zeta_i} \frac{f(\varphi(\zeta))}{\zeta\left(1 - z_k\,\overline{(\varphi\circ\rho)(\zeta)}\right)} \\ &= \sum_{i=1}^{m} f(\varphi(\zeta_i)) \operatorname*{Res}_{\zeta=\zeta_i} \frac{1}{\zeta\left(1 - z_k\,\overline{(\varphi\circ\rho)(\zeta)}\right)} \\ &= \sum_{j=1}^{n} f(z_j) \sum_{\varphi(\zeta_i)=z_j} \operatorname*{Res}_{\zeta=\zeta_i} \frac{1}{\zeta\left(1 - z_k\,\overline{(\varphi\circ\rho)(\zeta)}\right)} .\end{aligned}$$

We now use the definition of the matrix M stated in the proposition to obtain the identity

$$(C_\varphi^* C_\varphi f)(z_k) = \sum_{j=1}^{n} m_{jk}\, f(z_j). \tag{4}$$

We may use equation (4) to show explicitly how $C_\varphi^* C_\varphi$ acts on the kernel functions K_{z_k}, for $1 \le k \le n$:

$$\begin{aligned}(C_\varphi^* C_\varphi K_{z_k})(z) &= \langle C_\varphi^* C_\varphi K_{z_k}, K_z\rangle = \overline{\langle C_\varphi^* C_\varphi K_z, K_{z_k}\rangle} \\ &= \overline{(C_\varphi^* C_\varphi K_z)(z_k)} = \sum_{j=1}^{n} \overline{m_{jk}\, K_z(z_j)} = \sum_{j=1}^{n} \overline{m_{jk}}\, K_{z_j}(z).\end{aligned}$$

Let $W = \operatorname{Span}(\{K_{z_k}\}_{k=1}^{n})$. Since W is invariant under $C_\varphi^* C_\varphi$, we may use the same argument as in Theorem 2.4 to show that $\|C_\varphi^* C_\varphi\|$ is the greatest eigenvalue of the operator on W. Let $g \in W$ be an eigenfunction of $C_\varphi^* C_\varphi$, with $g = \sum_{k=1}^{n} c_k K_{z_k}$. Let $\mathbf{c} \in \mathbb{C}^n - \{\mathbf{0}\}$ be the vector with components $\{c_k\}_{k=1}^{n}$. Then, using our expression for $C_\varphi^* C_\varphi K_{z_k}$, we have $M^*\mathbf{c} = \lambda\,\mathbf{c}$. Hence $\lambda = \|C_\varphi\|^2$ is the greatest solution to the equation $|M^* - \lambda I| = 0$, where M^* is the conjugate transpose of M and I is the identity matrix. Since $\lambda \in \mathbb{R}$, this is equivalent to the equation $|M - \lambda I| = 0$. Therefore, $\|C_\varphi\|^2$ is the greatest eigenvalue of M. □

We now show how Theorem 6.1 can be used to provide a new proof for C. Cowen's formula ([2] or [3, p. 324]) for the norm of a composition operator with linear symbol.

Proposition 6.2 (Cowen)**.** *Let $\varphi(z) = sz + t$, with $|s| + |t| < 1$. Then*

$$\|C_\varphi\|^2 = \frac{2}{1 + |s|^2 - |t|^2 + \sqrt{(1 - |s|^2 + |t|^2)^2 - 4|t|^2}} . \tag{5}$$

Proof. Note that $\|\varphi\|_\infty < 1$, so C_φ is compact and hence norm-attaining. Let

$$\zeta_1 = \frac{1 - |s|^2 - |t|^2 - \sqrt{(1 - |s|^2 - |t|^2)^2 - 4|s|^2|t|^2}}{2\, s\, \bar{t}}$$

and let $z_1 = \varphi(\zeta_1)$ (so $A = \{\zeta_1\}$ and $B = \{z_1\}$). It is not too difficult to check that ζ_1 is the one and only root of $\zeta\left(1 - z_1\,\overline{(\varphi \circ \rho)(\zeta)}\right) = \zeta\left(1 - z_1\bar{t}\right) - z_1\bar{s}$. The condition that $\varphi(A) \subset B$ is true by the definition of z_1. We are now in a position to apply Theorem 6.1. The matrix M becomes a 1×1 matrix, with its only entry equal to

$$m_{11} = \frac{2}{1 + |s|^2 - |t|^2 + \sqrt{(1 - |s|^2 + |t|^2)^2 - 4|t|^2}}\,.$$

Hence the only eigenvalue of M is given by the expression above, so by Theorem 6.1, this is equal to $\|C_\varphi\|^2$. □

Although the above equation (5) also holds when $|s| + |t| = 1$, our methods fail in this case since ζ_1 falls on $\partial\mathbb{D}$, and, in fact, the operator C_φ is not norm-attaining.

The above proposition uses only the "$n = 1$" version of Theorem 6.1. The "$n = 2$" version of the theorem, with $B = \{0, \varphi(0)\}$, is just our earlier Theorem 2.4. For $n \geq 3$, it was pointed out by the referee for this paper that linear fractional examples can be found, as in Hammond's work [5, Section 7], by using

$$\varphi(z) = \frac{(r-1)\,z - (n-1)}{-nz + r}$$

for $r > n$. The operator $C_\varphi$then satisfies the hypotheses of Theorem 6.1, with $B = \{\varphi\,(0)\,, \tau\,(\varphi\,(0))\,, \tau\,(\tau\,(\varphi\,(0)))\,, \ldots, \tau_{n-1}\,(\varphi\,(0)) = 0\}$. More complicated examples for the $n \geq 3$ version of the theorem could surely be found, but they are beyond the scope of the current work.

References

[1] P.S. Bourdon., E.E. Fry, C. Hammond, and C.H. Spofford, *Norms of linear-fractional composition operators*, Trans. Amer. Math. Soc. 356 (2004) 2459–2480.

[2] C. Cowen *Linear fractional composition operators on H^2*, Integral Equations Operator Theory, 11 (1988), 151-160.

[3] C. Cowen and B. MacCluer, *Composition Operators on Spaces of Analytic Functions*, 1995: CRC Press, Boca Raton.

[4] C. Hammond, *On the norm of a composition operator*, Ph.D. Thesis, University of Virginia, 2003.

[5] C. Hammond, *On the norm of a composition operator with linear fractional symbol*, Acta Sci. Math. (Szeged) 69 (2003) 813–829.

[6] E. Nordgren, *Composition operators*, Canadian J. Math., 20 (1968), 442-449.

[7] J. H. Shapiro, *Composition Operators and Classical Function Theory*, 1993: Springer-Verlag, New York.

Princeton University, Princeton, NJ 08544
E-mail address: effinger@princeton.edu

University of California, Berkeley, CA 94720
E-mail address: asj@uclink.berkeley.edu

California Polytechnic State University, San Luis Obispo, CA 93407
E-mail address: jrreed@calpoly.edu

California Polytechnic State University, San Luis Obispo, CA 93407
E-mail address: jshapiro@calpoly.edu