Extremal non-compactness of composition operators with linear fractional symbol

Estelle L. Basor [*,1], Dylan Q. Retsek

Department of Mathematics, California Polytechnic State University, San Luis Obispo, CA 93407, USA

Received 27 July 2005

Available online 26 October 2005

Submitted by J.H. Shapiro

Abstract

We realize the norms of certain composition operators C_φ with linear fractional symbol acting on the Hardy space in terms of the roots of associated hypergeometric functions. This realization leads to simple necessary and sufficient conditions on φ for C_φ to exhibit extremal non-compactness, establishes equivalence of cohyponormality and cosubnormality of composition operators with linear fractional symbol, and yields a complete classification of those linear fractional φ that induce composition operators whose norms are determined by the action of the adjoint C_φ^* on the normalized reproducing kernels in H^2.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Composition operators; Extremal non-compactness; Cohyponormality

1. Introduction

Let $\mathbb{D} = \{z \in \mathbb{C}: |z| < 1\}$ be the unit disk in the complex plane. The classical Hardy space H^2 consists of those holomorphic functions on $\mathbb{D}$ whose Taylor coefficients about the origin are square summable. The space H^2 is a Hilbert space with inner product given by

$$\left\langle \sum_{j=0}^{\infty} a_j z^j, \sum_{k=0}^{\infty} b_k z^k \right\rangle = \sum_{j=0}^{\infty} a_j \bar{b}_j.$$

As a consequence of Cauchy's integral formula, point evaluation is a bounded linear functional on H^2. Thus, for each $w \in \mathbb{D}$, the Riesz representation theorem guarantees a unique function $K_w \in H^2$ such that

$$f(w) = \langle f, K_w \rangle$$

for all $f \in H^2$. The function K_w is called the reproducing kernel at the point w and is given by $K_w(z) = (1 - \bar{w}z)^{-1}$.

Given an analytic map $\varphi : \mathbb{D} \to \mathbb{D}$, the composition operator C_φ is defined on H^2 by $C_\varphi f = f \circ \varphi$. A direct application of Littlewood's subordination principle [13] shows that C_φ is in fact a bounded linear operator on H^2. Though boundedness is established with relative ease, in only special cases is $\|C_\varphi\|$ known. In general, calculating the norm of a composition operator appears to be a difficult problem.

Recall that a bounded linear operator T on a Hilbert space is *compact* if the image of the closed unit ball under T has compact closure. Thus, one way to try to compute $\|C_\varphi\|$ is to compare it to the essential norm

$$\|C_\varphi\|_e = \inf_{K \text{ cpt}} \|C_\varphi + K\|.$$

While it is clear that $\|C_\varphi\|_e \leqslant \|C_\varphi\|$, this inequality may be strict. For example, if $\varphi(z) = 2/(3 - z)$, then $0 < \|C_\varphi\|_e < \|C_\varphi\|$ (see [3, Theorem 3.9]). Because computing essential norms of composition operators on the Hardy space is often easier than computing norms, one would be interested to know precisely which self-maps φ of the unit disk result in composition operators having norm equal to essential norm. Such operators are said to exhibit *extremal non-compactness*.

In this paper we obtain necessary and sufficient conditions on the linear fractional map $\varphi : \mathbb{D} \to \mathbb{D}$ for the equality $\|C_\varphi\|_e = \|C_\varphi\|$ to hold. Our method involves interpreting the norm representation recently discovered by Bourdon et al. [3] in terms of hypergeometric function theory. In the next section we develop the background necessary for the sequel.

2. Background

For any holomorphic map $\varphi : \mathbb{D} \to \mathbb{D}$ we have

$$\frac{1}{1 - |\varphi(0)|^2} \leqslant \|C_\varphi\|^2 \leqslant \frac{1 + |\varphi(0)|}{1 - |\varphi(0)|} \tag{1}$$

(see, e.g., [9, Corollary 3.7]). Equality is achieved on the left when φ is constant and on the right when φ is inner [14]. The precise value of $\|C_\varphi\|$ is known in a few other cases as well. For example, when $\varphi(0) = 0$, Eq. (1) implies that $\|C_\varphi\| = 1$. Or when $\varphi(z) = sz + t$ (s and t necessarily satisfying $|s| + |t| \leqslant 1$), C. Cowen has computed

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

in [6].

In attempting to compute the norm of a linear operator $T : \mathcal{H} \to \mathcal{H}$, one often works with the positive operator T^*T and attempts to compute $\|T^*T\| = \|T\|^2$. The ingredients that can make this approach particularly successful when dealing with the composition operator $C_\varphi : H^2 \to H^2$ are twofold.

First, Hammond shows in [12] that if φ is an analytic self-map of $\mathbb{D}$ such that $\|C_\varphi\| > \|C_\varphi\|_e$, then there exists a nowhere vanishing function $g \in H^2$ such that $C_\varphi^* C_\varphi g = \|C_\varphi\|^2 g$.

Second, when $\varphi(z) = (az + b)/(cz + d)$ is a non-constant linear fractional map the adjoint of C_φ is computable via Cowen's formula

$$C_\varphi^* = T_\gamma C_\sigma T_\nu^*$$

where $\gamma(z) = 1/(-\bar{b}z + \bar{d})$, $\nu(z) = cz + d$, $\sigma(z) = (\bar{a}z - \bar{c})/(-\bar{b}z + \bar{d})$ and T_h is the analytic Toeplitz operator with symbol h.

Using these two facts, the quest to compute $\|C_\varphi\|$ when $\varphi(z) = (az + b)/(cz + d)$ and $\|C_\varphi\| > \|C_\varphi\|_e$ is reduced to looking for the largest eigenvalue of $C_\varphi^* C_\varphi$. Toward this end, Hammond established in [12] that if f is an eigenvector for $C_\varphi^* C_\varphi$ with corresponding eigenvalue λ, the functional equation

$$\lambda f(z) = \psi(z) f\big(\tau(z)\big) + \chi(z)\lambda f(0) \tag{2}$$

holds at every z for which the auxiliary maps

$$\psi(z) = \frac{(\overline{ad} - \overline{bc})z}{(\bar{a}z - \bar{c})(-\bar{b}z + \bar{d})}, \quad \chi(z) = \frac{\bar{c}}{-\bar{a}z + \bar{c}}, \quad \text{and} \quad \tau(z) = \varphi\big(\sigma(z)\big) \tag{3}$$

are defined.

Iteration of Eq. (2) ultimately leads to the following theorem of Bourdon, Fry, Hammond and Spofford which appears as Theorem 3.5 and Corollary 3.6 in [3].

Theorem 2.1. *Let $\varphi(z) = (az + b)/(cz + d)$ be a non-automorphic linear fractional mapping that fixes the point* 1. *Then $\|C_\varphi\| > \|C_\varphi\|_e$ if and only if there exists a number $\Lambda > \|C_\varphi\|_e^2$ such that*

$$\sum_{k=0}^\infty \chi\big(\tau^{[k]}(\varphi(0))\big) \left[\prod_{m=0}^{k-1} \psi\big(\tau^{[m]}(\varphi(0))\big) \right] \left(\frac{1}{\Lambda} \right)^{k+1} = 1. \tag{4}$$

Moreover, the largest Λ (when there is one) for which Eq. (4) holds is $\|C_\varphi\|^2$.

In general, it is not easy to tell whether Eq. (4) has any solutions. In special circumstances, however, one can use this equation to good effect. For instance, in [3] Bourdon et al. show that if $\varphi(z) = (\alpha - 1)/(\alpha - z)$ for some $\alpha > 1$, then $\|C_\varphi\| > \|C_\varphi\|_e$ (the specific case $\alpha = 3$ was mentioned in the Introduction). The proof is via the intermediate value theorem and depends critically upon the fact that the coefficients of the series in (4) are all real for this particular kind of φ.

It is a major goal of this paper to circumvent this dependence on real coefficients in the series of Theorem 2.1 and show that $\|C_\varphi\|_e$ very rarely equals $\|C_\varphi\|$ for linear fractional φ. The key observation toward this end is the subject of the next section.

3. Hypergeometric series

A hypergeometric series is a series $\sum a_k$ such that the ratio a_{k+1}/a_k is a rational function of k. In general, the numerator and denominator of a_{k+1}/a_k will be polynomials in k of arbitrary degree. Of particular importance to us, however, is the special case

$$_2F_1(a, b; c; z) \equiv \sum_{k=0}^\infty \frac{(a)_k (b)_k}{(c)_k k!} z^k, \tag{5}$$

where $(\alpha)_0 = 1$ and $(\alpha)_k = (\alpha)(\alpha + 1) \cdots (\alpha + k - 1)$ for $k = 1, 2, 3, \ldots$.

The function $_2F_1(a, b; c; z)$ is called the hypergeometric series of variable z with parameters a, b, c. Note that the parameter c cannot be a negative integer. By the ratio test, the series converges absolutely for $|z| < 1$ and therefore defines an analytic function on $\mathbb{D}$.

Thanks to Euler, Gauss, Riemann and others much is known about the function $_2F_1$, as well as generalized hypergeometric functions having more parameters. We shall need various facts concerning $_2F_1$ along the way, but for a full treatment see the Bateman manuscript [11] or the treatise of Andrews et al. [1].

The remainder of this section is devoted to showing that the series in (4) is in fact hypergeometric whenever φ is a non-automorphic, non-affine linear fractional self-map of $\mathbb{D}$ (extremal non-compactness of C_φ is already settled for automorphic and affine φ; see [3,6,16,17]). We begin with several lemmas.

Lemma 3.1. *Suppose $\varphi : \mathbb{D} \to \mathbb{D}$ is a non-affine linear fractional map that fixes the point* 1. *Then φ is of the form*

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}, \tag{6}$$

where $q = \varphi'(1) > 0$ and $d \in \mathbb{C} \setminus \bar{\mathbb{D}}$.

Proof. Because φ is not affine, the linear coefficient in the denominator is non-zero. We may therefore assume without loss of generality that $\varphi(z) = (az + b)/(z + d)$. Since $\varphi(1) = 1$, $a + b = 1 + d$. Thus,

$$\varphi(z) = \frac{az + (1 + d - a)}{z + d}.$$

Differentiating, we find that

$$\varphi'(z) = \frac{(a - 1)(d + 1)}{(z + d)^2}. \tag{7}$$

Now, because $\varphi(1) = 1$ the Julia–Carathéodory theorem implies that $q \equiv \varphi'(1) > 0$. Substituting $z = 1$ in (7) yields $a = 1 + q + qd$ and

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}$$

as claimed. $\square$

We note that not all maps of the form in the above lemma are self-maps of the disk. Indeed, $\varphi(0) = 1 - q(d + 1)/d$ shows that q and d cannot be chosen independently if one hopes to obtain a self-map. Examination of the conformally equivalent map $f = T \circ \varphi \circ T^{-1}$, where $T(z) = (1 + z)/(1 - z)$ maps the disk onto the right half-plane, shows that $\varphi : \mathbb{D} \to \mathbb{D}$ if and only if $\Re\{(d - 1)/(d + 1)\} \geqslant q$. This condition will play a role later.

Lemma 3.1 is useful in that it pares down the number of parameters we must deal with. Because we wish to eventually analyze the series of Theorem 2.1, we would also be well served to simplify the auxiliary maps associated to the linear fractional map φ. Recall that if $\varphi(z) = (az + b)/(cz + d)$, then $\sigma(z) = (\bar{a}z - \bar{c})/(-\bar{b}z + \bar{d})$ and $\tau(z) = \varphi(\sigma(z))$. The next lemma gives an explicit representation of τ in terms of the parameters q and d of Lemma 3.1.

Lemma 3.2. *Suppose* $\varphi : \mathbb{D} \to \mathbb{D}$ *is of the form*

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}.$$

Then the associated map τ *is of the form*

$$\tau^{[k]}(z) = T^{-1}\big(f^{[k]}\big(T(z)\big)\big) = \frac{(1 - kb)z + kb}{-kbz + 1 + kb}, \tag{8}$$

where

$$b = \frac{|d|^2 - q|1 + d|^2 - 1}{q|1 + d|^2}. \tag{9}$$

Proof. It is customary (and exceedingly useful) to work with the "right half-plane version" of τ. Specifically, let $f = T \circ \tau \circ T^{-1}$ where $T(z) = (1 + z)/(1 - z)$ is the conformal mapping that takes the unit disk onto the right half-plane. Note that $f = (T \circ \varphi \circ T^{-1}) \circ (T \circ \sigma \circ T^{-1})$ is just the composition of the right half-plane versions of φ and σ. We easily compute

$$\big(T \circ \varphi \circ T^{-1}\big)(z) = \frac{z}{q} + \frac{d - 1}{q(d + 1)} - 1$$

and

$$\big(T \circ \sigma \circ T^{-1}\big)(z) = qz + \frac{\bar{d} - 1}{\bar{d} + 1} - q.$$

Composing, we obtain

$$f(z) = \big(T \circ \varphi \circ T^{-1}\big) \circ \big(T \circ \sigma \circ T^{-1}\big)(z) = z + \frac{2(|d|^2 - q|1 + d|^2 - 1)}{q|1 + d|^2} \equiv z + 2b.$$

The utility of this right half-plane version of τ is now evident; the map $f(z)$ is just translation of the half-plane by the non-negative real number $2b$. The k-fold iterate $f^{[k]}$ is therefore given by $f^{[k]}(z) = z + k(2b)$ and we finally obtain

$$\tau^{[k]}(z) = T^{-1}\big(f^{[k]}\big(T(z)\big)\big) = \frac{(1 - kb)z + kb}{-kbz + 1 + kb} \tag{10}$$

as desired. □

With this firm grip on the auxiliary map τ, we are now in a position to show that the series

$$\sum_{k=0}^{\infty} \chi\big(\tau^{[k]}\big(\varphi(0)\big)\big)\left[\prod_{m=0}^{k-1} \psi\big(\tau^{[m]}\big(\varphi(0)\big)\big)\right] x^{k+1}$$

is hypergeometric whenever φ is a non-affine linear fractional self map of the unit disk. For convenience we remind the reader of the auxiliary maps corresponding to $\varphi(z) = (az + b)/(cz + d)$ that were discussed in Section 2:

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

Theorem 3.3. *Let $\varphi : \mathbb{D} \to \mathbb{D}$ be a non-affine, non-automorphic linear fractional map that fixes the point* 1. *Then*

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}$$

and

$$\sum_{k=0}^{\infty} \chi\left(\tau^{[k]}(\varphi(0))\right) \left[\prod_{m=0}^{k-1} \psi\left(\tau^{[m]}(\varphi(0))\right) \right] x^{k+1} = 1 - {}_2F_1(\alpha, \beta; \delta; x/q) \qquad (11)$$

where

$$\alpha = \frac{1 + d}{qb|1 + d|^2}, \quad \beta = \frac{d + d\bar{d}}{qb|1 + d|^2}, \quad and \quad \delta = \frac{1}{qb}.$$

Proof. The form of φ is guaranteed by Lemma 3.1. Thus, $\varphi(0) = (d - q - qd)/d$ and Lemma 3.2 shows that

$$\tau^{[k]}(\varphi(0)) = \frac{(1 - kb)\varphi(0) + kb}{-kb\varphi(0) + 1 + kb} = 1 - \frac{q(1 + d)}{d + qb(1 + d)k}$$

(where b is given by Eq. (9) above).

Let $\sum_{k=0}^{\infty} a_k x^{k+1}$ be the series on the left side of equality (11). To show that the series is hypergeometric, we analyze the quotient

$$\frac{a_{k+1}}{a_k} = \frac{\chi(\tau^{[k+1]}(\varphi(0)))\psi(\tau^{[k]}(\varphi(0)))}{\chi(\tau^{[k]}(\varphi(0)))}. \qquad (12)$$

This appears to be a daunting task. Upon substituting the formulas for the functions ψ and χ (with the coefficients for φ already in terms of q and d) and the value for $\tau^{[k]}(\varphi(0))$ obtained above, a fair amount of simplification takes place. Continuing from (12) we have

$$\frac{a_{k+1}}{a_k} = \left\{ \frac{q[d(1 + \bar{d}) - q|1 + d|^2 + qb|1 + d|^2 k]}{[q(|d|^2 - 1) + q^2(b - 1)|1 + d|^2 + q^2 b|1 + d|^2 k]} \right\}$$

$$\times \left\{ \frac{d(1 + \bar{d}) + qb|1 + d|^2 + qb|1 + d|^2 k}{[q(|d|^2 - 1 + |1 + d|^2) - q^2|1 + d|^2 + q^2 b|1 + d|^2 k]} \right\}.$$

While we are encouraged because a_{k+1}/a_k indeed appears to be a rational function of k, this current incarnation of the quotient is far from practical. Factoring out a q from each term in the denominator and dividing top and bottom by $(qb|1 + d|^2)^2$ we finally arrive at the more palatable

$$\frac{a_{k+1}}{a_k} = \frac{1}{q} \frac{[\alpha + k + 1][\beta + k + 1]}{[k + 2][\delta + k + 1]}, \qquad (13)$$

where

$$\alpha = \frac{1 + d}{qb|1 + d|^2}, \quad \beta = \frac{d + d\bar{d}}{qb|1 + d|^2}, \quad and \quad \delta = \frac{1}{qb}$$

are just as in the statement of the theorem.

It remains to prove that $\sum_{k=0}^{\infty} a_k x^{k+1} = 1 - {}_2F_1(\alpha, \beta; \delta; x/q)$. Consider the series $\sum_{k=0}^{\infty} b_k x^{k+1}$ where the coefficient b_k is given by

$$b_k = \frac{(\alpha + 1)_k (\beta + 1)_k}{(\delta + 1)_k (k + 1)!} \left(\frac{1}{q} \right)^k.$$

By design, the quotient $b_{k+1}/b_k = a_{k+1}/a_k$ satisfies Eq. (13). Because the first coefficient of any such series determines the remaining coefficients, we must have

$$\sum_{k=0}^{\infty} a_k x^{k+1} = \frac{a_0}{b_0} \sum_{k=0}^{\infty} b_k x^{k+1} = \frac{a_0}{b_0} \sum_{k=0}^{\infty} \frac{(\alpha+1)_k (\beta+1)_k}{(\delta+1)_k (k+1)!} \left(\frac{1}{q}\right)^k x^{k+1}. \tag{14}$$

Now, utilizing the facts that $b_0 = 1$, $a_0 = \chi(\varphi(0))$, and $\zeta(\zeta+1)_k = (\zeta)_{k+1}$ whenever $\zeta \neq 0$ and $k = 0, 1, 2, \dots$ (recall that α, β and δ are all non-zero), we continue from (14) to obtain

$$q \frac{\chi(\varphi(0))\delta}{\alpha\beta} \sum_{k=0}^{\infty} \frac{(\alpha)_{k+1}(\beta)_{k+1}}{(\delta)_{k+1}(k+1)!} \left(\frac{x}{q}\right)^{k+1} = q \frac{\chi(\varphi(0))\delta}{\alpha\beta} \big[{}_2F_1(\alpha, \beta; \delta; x/q) - 1 \big].$$

Finally, a simple computation shows that

$$\frac{\chi(\varphi(0))\delta}{\alpha\beta} = -\frac{1}{q}$$

and the proof is complete. $\quad\square$

It is worth noting that the maps φ of Theorem 3.3 above may be written in terms of the parameters α and β from the outset. Indeed, if $\varphi : \mathbb{D} \to \mathbb{D}$ is a non-automorphic linear fractional map that fixes the point 1, then φ may be written

$$\varphi(z) = \frac{(\bar{\beta}-1)z + \alpha + 1}{\bar{\alpha}z + \beta},$$

where α, β and $\delta = \bar{\alpha} + \beta$ are precisely the parameters that appear in the associated hypergeometric series. While certainly appealing on aesthetic grounds, we will see that the parameter d has a good deal to say about two interesting operator theoretic questions regarding C_φ. Since one may determine d at a glance given an arbitrary non-affine linear fractional self-map of the disk, it is convenient to continue working with our qd-parameterization.

Because so much is known about hypergeometric functions, we may parlay the information contained in Theorem 3.3 to garner information about the norms of a great many composition operators with linear fractional symbol. Specifically, the work of the next section will show that such operators are rarely extremally non-compact and, as a consequence, their norms are rarely exhibited by the action of the adjoint C_φ^* acting on the normalized reproducing kernels in H^2.

4. Extremal non-compactness

Given a linear fractional map $\varphi : \mathbb{D} \to \mathbb{D}$, we would like to know the exact conditions under which $\|C_\varphi\| = \|C_\varphi\|_e$. Because $\|\varphi\|_\infty < 1$ implies that C_φ is compact, we limit our discussion to φ for which $\|\varphi\|_\infty = 1$. In this case, either φ is an automorphism or else $\varphi(\mathbb{D})$ is a disk internally tangent to $\partial\mathbb{D}$ at a single point. If φ is an automorphism, then $\|C_\varphi\| = \|C_\varphi\|_e$ (see, e.g., [17]). We may therefore further limit our discussion to maps φ such that there exists a single pair of points ζ and η on $\partial\mathbb{D}$ such that $\varphi(\zeta) = \eta$.

Having gone this far, we note that without loss of generality we may still further assume that $\zeta = \eta = 1$. To wit, Bourdon et al. note in [3] that for any unimodular constants ζ and η, the operators $C_{\zeta z}$ and $C_{\bar{\eta}z}$ are unitary and thus $\|C_\varphi\| = \|C_\varphi\|_e$ if and only if $\|C_{\zeta z} C_\varphi C_{\bar{\eta}z}\| = \|C_{\zeta z} C_\varphi C_{\bar{\eta}z}\|_e$. But when $\varphi(\zeta) = \eta$, $\bar{\eta}\varphi(\zeta z)$ is a linear fractional self-map of the disk that fixes the point 1. It therefore suffices to know how to compare $\|C_\varphi\|$ and $\|C_\varphi\|_e$ when $\varphi(1) = 1$.

Finally, affine maps $\varphi(z) = sz + t$ satisfying $\|\varphi\|_\infty = 1$ also satisfy $\|C_\varphi\| = \|C_\varphi\|_e$; Cowen's norm formula in [6] and Shapiro's essential norm formula in [16] may be seen to coincide in this case. Thus, we will assume throughout this section that $\varphi : \mathbb{D} \to \mathbb{D}$ *is a non-affine, non-automorphic linear fractional map that fixes the point* 1. In the face of the rather restrictive hypotheses of Theorem 3.3, this assumption lends credibility to the work of Section 3 and, via Lemma 3.1, allows us to use the representation

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}$$

freely. We begin with a lemma.

Lemma 4.1. *Suppose* $\varphi : \mathbb{D} \to \mathbb{D}$ *is given by*

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}$$

and the number

$$\alpha = \frac{1 + d}{qb|1 + d|^2} \leqslant -1.$$

Then $\|C_\varphi\| > \|C_\varphi\|_e$.

Proof. Say $\alpha = -t$ for some $t \geqslant 1$. Then $d < -1$ (recall that q and b are both positive) and substituting the value of b given by (9) in terms of q and d we have

$$-t = \frac{1}{qb(1 + d)} = \frac{1 + d}{d^2 - q(1 + d)^2 - 1}.$$

Solving for q, we obtain $q = (1 + t(d - 1))/(td + t)$. Because φ is non-automorphic and fixes 1, Shapiro's essential norm formula [16] shows that $\|C_\varphi\|_e^2 = 1/q$. Thus, since $d + 1 \leqslant (d + 1)/t < 0$,

$$\|C_\varphi\|_e^2 = \frac{td + t}{1 + t(d - 1)} \leqslant \frac{1 + d}{d} = 1 + \frac{1}{d} < 1 \leqslant \|C_\varphi\|^2,$$

where the last inequality above is true for general self-maps φ by Eq. (1). $\square$

The next proposition shows that the conclusion of the previous lemma holds in general when d is negative.

Proposition 4.2. *Suppose* $\varphi : \mathbb{D} \to \mathbb{D}$ *is given by*

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}.$$

If $d < -1$, *then* $\|C_\varphi\| > \|C_\varphi\|_e$.

Proof. By Lemma 4.1 we may assume that $-1 < \alpha < 0$. According to Theorems 2.1 and 3.3, $\|C_\varphi\| > \|C_\varphi\|_e$ if and only if the equation $_2F_1(\alpha, \beta; \delta; x/q) = 0$ has a solution in the interval $(0, q)$. Thus, we aim to prove that $_2F_1(\alpha, \beta; \delta; x) = 0$ has a root in $(0, 1)$ where

$$\alpha = \frac{1 + d}{qb|1 + d|^2}, \quad \beta = \frac{d + d\bar{d}}{qb|1 + d|^2}, \quad \text{and} \quad \delta = \frac{1}{qb}.$$

When the number d is real there are some noteworthy relations among the constants α, β, and δ. First, $\beta = \alpha d$ when $d \in \mathbb{R}$; since α and d have the same sign (again, recall that b and q are positive), we conclude that $\beta > 0$ whenever d is real. Second, we in general have

$$\delta - \beta = \frac{1}{qb} - \frac{d + d\bar{d}}{qb|1+d|^2} = \frac{1 + \bar{d}}{qb|1+d|^2} = \bar{\alpha}.$$

So when d is real, $\delta - \beta = \alpha$.

In general, the series $_2F_1(\alpha, \beta; \delta; x)$ has radius of convergence $R = 1$ and diverges at $x = 1$ when $\Re\{\delta - \alpha - \beta\} \leqslant 0$ (except in the case $\alpha = -n$, which we are not currently considering). In the present case where $\delta = \alpha + \beta$, the behavior of the series as $x \to 1^-$ was determined by Gauss (see [1, Theorem 2.1.3]):

$$\lim_{x \to 1^-} \frac{_2F_1(\alpha, \beta; \delta; x)}{-\log(1-x)} = \frac{\Gamma(\delta)}{\Gamma(\alpha)\Gamma(\beta)}. \tag{15}$$

Here $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$ is the gamma function familiar from complex analysis, analytically continued to the plane minus zero and the negative integers.

Since $-1 < \alpha < 0$, $\beta > 0$, and $\delta > 0$, the right-hand side of (15) is a negative real number. Because $-\log(1-x) \to \infty$ as $x \to 1^-$, we conclude that $_2F_1(\alpha, \beta; \delta; x) \to -\infty$ as $x \to 1^-$. The function $_2F_1(\alpha, \beta; \delta; x)$ is continuous on $0 \leqslant x < 1$ and $_2F_1(\alpha, \beta; \delta; 0) = 1$, so the intermediate value theorem implies that $_2F_1(\alpha, \beta; \delta; x)$ has a root in $(0, 1)$. $\square$

In spirit, the analysis of the series $_2F_1$ in the proposition above is very much akin to that found in [3, Theorem 3.7]. In fact, the astute reader may have already noticed that the map φ of Proposition 4.2 is indeed the member $\phi_{r,s}$ of the Cowen–Kriete family studied there with $r = (d - q - qd)^{-1}$ and $s = q$ (about which we shall have more to say later). In this sense, the hypergeometric approach simply offers an alternative light in which to cast these types of series equations.

The greatest strength of the hypergeometric approach, however, is its capacity to give information when the coefficients of the linear fractional map φ are complex-valued. One readily sees that the proof of the preceding proposition fails (on several counts) when d is not real. The remedy relies on a lovely transformation of Pfaff (see [1, Theorem 2.2.5], for a proof).

Theorem 4.3 *(Pfaff's transformation).*

$$_2F_1(a, b; c; x) = (1-x)^{-a}\, _2F_1\left(a, c - b; c; \frac{x}{x-1}\right). \tag{16}$$

We remark that one must take care in interpreting the right side of (16). For $0 < x < 1$ the argument $x/(x-1)$ takes values in $(-\infty, 0)$, yet the series definition of $_2F_1$ is valid only in the unit disk. There are, however, various integral representations of hypergeometric functions that allow analytic continuation to the whole plane minus the ray $[1, \infty)$. Interpreted this way, Pfaff's transformation is valid and we henceforth let context dictate what we mean by the function $_2F_1$. For a much more extensive treatment of analytic continuation of hypergeometric functions refer to [1] or [11].

We are now in a position to prove the main theorem of this section.

Theorem 4.4. *Suppose $\varphi : \mathbb{D} \to \mathbb{D}$ is given by*

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}.$$

Then $\|C_\varphi\| = \|C_\varphi\|_e$ *if and only if* $d > 1$.

Proof. We know already by Theorems 2.1 and 3.3 that $\|C_\varphi\| > \|C_\varphi\|_e$ if and only if the series $_2F_1(\alpha, \beta; \delta; x)$ has a root in $(0, 1)$.

Suppose $d > 0$. The coefficients of the series $_2F_1(\alpha, \beta; \delta; x)$ are then clearly positive. Since $_2F_1(\alpha, \beta; \delta; 0) = 1$, the series has no roots in $(0, 1)$ and therefore $\|C_\varphi\| = \|C_\varphi\|_e$.

For the converse, suppose that d is not positive. Proposition 4.2 handles the case $d < -1$, so we are left to consider whether $_2F_1(\alpha, \beta; \delta; x) = 0$ holds for any x in $(0, 1)$ when $0 < |\arg d| < \pi$. Because $\delta - \beta = \bar{\alpha}$, Pfaff's transformation implies

$$_2F_1(\alpha, \beta; \delta; x) = (1 - x)^{-\alpha} \, _2F_1\left(\alpha, \bar{\alpha}; \delta; \frac{x}{x - 1}\right). \tag{17}$$

The utility of Eq. (17) is that the hypergeometric series on the right side is *real* for all $x \in (0, 1)$. This is easy to see when $0 < x < 1/2$; in this case, $x/(x - 1)$ is in the unit disk and the coefficients in the series definition of $_2F_1$ are all real. When $1/2 \leqslant x < 1$, things are less clear. For these values of x, the argument $x/(x - 1)$ lies in the interval $(-\infty, -1]$ and we must take $_2F_1$ to mean the analytic continuation of the series valid in the disk. Letting $t = x/(x - 1)$ in this case, the identity

$$\frac{_2F_1(\alpha, \bar{\alpha}; \delta; t)}{\Gamma(\delta)} = \frac{\Gamma(\bar{\alpha} - \alpha)}{\Gamma(\bar{\alpha})\Gamma(\delta - \alpha)} (-t)^{-\alpha} \, _2F_1(\alpha, 1 - \delta + \alpha; 1 - \bar{\alpha} + \alpha; 1/t)$$
$$+ \frac{\Gamma(\alpha - \bar{\alpha})}{\Gamma(\alpha)\Gamma(\delta - \bar{\alpha})} (-t)^{-\bar{\alpha}} \, _2F_1(\bar{\alpha}, 1 - \delta + \alpha; 1 - \alpha + \bar{\alpha}; 1/t), \tag{18}$$

which follows from Barnes' integral representation for $_2F_1$ and is valid provided $\alpha - \bar{\alpha}$ is not an integer (see [11, Section 2.1.4]), shows that

$$\overline{_2F_1(\alpha, \bar{\alpha}; \delta; x/(x - 1))} = \,_2F_1\bigl(\alpha, \bar{\alpha}; \delta; x/(x - 1)\bigr)$$

holds for $1/2 \leqslant x < 1$ (due to the symmetry in α and $\bar{\alpha}$).

The point here is that $(1 - x)^{-\alpha}$ is non-zero for $0 < x < 1$, so according to Eq. (17) $_2F_1(\alpha, \beta; \delta; x)$ has a root in $(0, 1)$ if and only if $_2F_1(\alpha, \bar{\alpha}; \delta; t)$ has a root t in $(-\infty, 0)$. Since we have established that $_2F_1(\alpha, \bar{\alpha}; \delta; t)$ is real for these values of t, we may once again employ the intermediate value theorem to guarantee the desired root. Specifically, it now suffices to show that $_2F_1(\alpha, \bar{\alpha}; \delta; t) < 0$ for some $t \in (-\infty, 0)$.

Toward this goal, observe that Eq. (18) may be written

$$_2F_1(\alpha, \bar{\alpha}; \delta; t) = \frac{\Gamma(\bar{\alpha} - \alpha)\Gamma(\delta)}{\Gamma(\bar{\alpha})\Gamma(\delta - \alpha)(-t)^\alpha}\left[1 + O\left(\frac{1}{t}\right)\right]$$
$$+ \frac{\Gamma(\alpha - \bar{\alpha})\Gamma(\delta)}{\Gamma(\alpha)\Gamma(\delta - \bar{\alpha})(-t)^{\bar{\alpha}}}\left[1 + O\left(\frac{1}{t}\right)\right] \quad (t \to -\infty).$$

Letting $s = -t$, it therefore suffices to show that

$$(A + Bi)s^{-\alpha} + (A - Bi)s^{-\bar{\alpha}} < 0$$

for arbitrarily large positive values of s and

$$A + Bi \equiv \frac{\Gamma(\bar{\alpha} - \alpha)\Gamma(\delta)}{\Gamma(\bar{\alpha})\Gamma(\delta - \alpha)}.$$

Now, if $\alpha = x + iy$, then

$$(A + Bi)s^{-\alpha} + (A - Bi)s^{-\bar{\alpha}} = 2s^{-x}\big[A\cos\big(y\ln(s)\big) + B\sin\big(y\ln(s)\big)\big].$$

Note that A and B cannot both be zero since the function $_2F_1(\alpha, \bar{\alpha}; \delta; t)$ is not identically zero for $t < 0$. Also note that because $0 < |\arg d| < \pi$, α is not real. In particular, $y \neq 0$ and, consequently, the function

$$G(\omega) = A\cos(y\omega) + B\sin(y\omega)$$

is negative for arbitrarily large positive values of ω. For instance, if $A > 0$ then $G((2k+1)\pi/y) = -A$ for each $k \in \mathbb{N}$; the possibilities $A < 0$, $B > 0$, and $B < 0$ are handled in a similar fashion. $\quad\square$

This completes the classification of the extremely non-compact composition operators induced by linear fractional self maps of the disk. To summarize, there are three possibilities for a linear fractional map $\varphi : \mathbb{D} \to \mathbb{D}$:

- $\|\varphi\|_\infty < 1$, in which case $\|C_\varphi\| > \|C_\varphi\|_e = 0$.
- φ is an automorphism, in which case

$$\|C_\varphi\| = \|C_\varphi\|_e = \big[\big(1 + |\varphi(0)|\big)\big/\big(1 - |\varphi(0)|\big)\big]^{1/2}.$$

- The image $\varphi(\mathbb{D})$ is a disk internally tangent to $\partial\mathbb{D}$ at a single point. In this case, either $\varphi(z) = sz + t$ and $\|C_\varphi\| = \|C_\varphi\|_e$ or else φ is a non-affine, non-automorphic map. If the latter is the case, then φ may be pre- and post-composed with rotations to produce a non-affine, non-automorphic map $\tilde{\varphi}$ that fixes the point 1 and induces a composition operator with the same norm and essential norm as C_φ. This new map must be of the form

$$\tilde{\varphi}(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}$$

and has $\|C_{\tilde{\varphi}}\| = \|C_{\tilde{\varphi}}\|_e$ if and only if $d > 1$.

Christopher Hammond has observed that this classification settles another closely related operator-theoretic question. This is the subject of the next section.

5. Cohyponormality

Recall that for a self-map φ, the Denjoy–Wolff point of φ is the unique point $a \in \overline{\mathbb{D}}$ such that $|\varphi'(a)| \leqslant 1$. If $|a| < 1$, then the spectral radius $r(C_\varphi)$ of C_φ is 1; if $|a| = 1$, then $r(C_\varphi) = [\varphi'(a)]^{-1/2}$ (see [9, Theorem 3.9]). When working with norms, it is thus sometimes helpful to deal with composition operators that enjoy equality of norm and spectral radius. The composition operators exhibiting some form of normality form just such a class (recall that $T : \mathcal{H} \to \mathcal{H}$ is normal if $TT^* = T^*T$).

A Hilbert space operator $T : \mathcal{H} \to \mathcal{H}$ is called *hyponormal* if $T^*T - TT^* \geqslant 0$ (equivalently, $\|Th\| \geqslant \|T^*h\|$ for all $h \in \mathcal{H}$). The operator T is *subnormal* if there exists an auxiliary Hilbert space $\mathcal{K} \supseteq \mathcal{H}$ and a normal operator $N : \mathcal{K} \to \mathcal{K}$ such that T is the restriction of N to $\mathcal{H}$. In this section we shall be concerned with deciding when the adjoint T^* satisfies these definitions, in which case the operator T is said to be *co*hyponormal or *co*subnormal.

The operator C_φ itself is known to be subnormal in just a few cases. These are: when $\varphi(z) = cz$ (these are precisely the normal C_φ), when φ is inner and fixes the origin (these are precisely the isometric C_φ), and when φ is linear fractional and C_φ is unitarily equivalent to $1 \oplus C_\psi^*$ with C_ψ^* subnormal.

Perhaps the best theorem on subnormality of C_φ^* is the following result of C. Cowen and T. Kriete in their article [7].

Theorem 5.1. *For $0 < s < 1$ and $0 \leqslant r \leqslant 1$, if*

$$\varphi_{r,s}(z) = \frac{(r+s)z + (1-s)}{r(1-s)z + (1+sr)},$$

then $C_{\varphi_{r,s}}^*$ is subnormal.

A nice discussion of the proof may be found in [9, pp. 314–317]. Under some additional smoothness hypotheses on φ, the converse to Theorem 5.1 is true as well [7]. The authors there suggest that the smoothness assumption serves only to facilitate the proof of the converse and that probably it is true without any additional requirements on φ.

Using the definitions above, it is not hard to show that any subnormal operator is hyponormal. Though not true in general, for adjoints of composition operators with linear fractional symbol the converse holds as well.

Theorem 5.2. *Let $\varphi : \mathbb{D} \to \mathbb{D}$ be linear fractional. If C_φ is cohyponormal, then C_φ is cosubnormal.*

Proof. Suppose C_φ^* is hyponormal. If C_φ^* is normal, then it is certainly subnormal. We may therefore assume without loss of generality that C_φ^*, hence C_φ, is not normal.

Theorem 8.4 of [9] then implies that φ has Denjoy–Wolff point $a \in \partial\mathbb{D}$ with $\varphi'(a) < 1$. Since both hypo- and subnormality are unitary invariants, we may assume (as usual) that $a = 1$. The remainder of the proof proceeds in cases.

If φ is affine, then $\varphi(z) = sz + (1-s)$ where $0 < s = \varphi'(1) < 1$. In this case, $\varphi = \varphi_{0,s}$ and C_φ^* is subnormal by Theorem 5.1.

If φ is an automorphism, then a priori it is either elliptic, parabolic, or hyperbolic. Because $\varphi(1) = 1$, φ is not elliptic. In addition, φ is not parabolic because in this case we have $r(C_\varphi) = 1 < \|C_\varphi\|$ by [9, Theorem 7.5] (recall that our cohyponormality assumption guarantees norm and spectral radius coincide). It is possible, however, for φ to be hyperbolic.

Suppose that $\varphi(z) = \lambda(w - z)/(1 - \bar{w}z)$ is hyperbolic. In this case, Theorem 7.4 of [9] asserts that $r(C_\varphi) = (\varphi'(1))^{-1/2}$. Equating this spectral radius to $\|C_\varphi\|$ forces $-1 < w < 0$ and one finds that in actuality $\varphi = \varphi_{r,s}$ where $r = 1$ and $s = \|C_\varphi\|^{-2}$. Thus, when φ is hyperbolic, Theorem 5.1 again implies that C_φ^* is subnormal.

Finally, if $\varphi : \mathbb{D} \to \mathbb{D}$ is a non-affine, non-automorphic linear fractional map that fixes the point 1, we are squarely in the realm of Theorem 4.4. That result, combined with the fact that essential norm equals spectral radius for a composition operator with non-automorphic linear fractional symbol having Denjoy–Wolff point on $\partial\mathbb{D}$ (see [3, Section 2] for the details), ensures that φ is of the form

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}$$

for some $d > 1$ and $0 < q = \varphi'(1) < 1$. Set $r = (d - q - qd)^{-1}$ and $s = q$. Since $\Re\{(d-1)/(d+1)\} \geqslant q$ is necessary for φ to map the disk into itself, we obtain $0 < r \leqslant 1$ and $0 < s < 1$. Thus, $\varphi = \varphi_{r,s}$ is yet again a member of the Cowen–Kriete family of Theorem 5.1 and consequently C_φ^* is subnormal in this final case as well. $\quad\square$

We remark that the family of self-maps φ with "qd-representation"

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d},$$

where $q > 0$ and $|d| > 1$ is distinct from the "extended" Cowen–Kriete family

$$\varphi_{r,s}(z) = \frac{(r+s)z + (1-s)}{r(1-s)z + (1+sr)},$$

where $0 \leqslant |r| \leqslant 1$ and $0 < s < 1$.

Lemma 3.1 shows that the non-affine members of the extended Cowen–Kriete family certainly have qd-representation, but the self-map $\varphi(z) = 1/(2 - z)$, which has $q = 1$ and $d = -2$, is not equal to $\varphi_{r,s}$ for any r and s. What has been most important to us, however, is that self-maps with qd-representation, $d > 1$, *do* belong the Cowen–Kriete family and consequently enjoy the attendant nice properties of maps therein.

We should also point out that Theorem 5.2 marks the culmination of steady progress in that direction. Earlier relevant work includes [5,8,10]. Subsuming the more recent completion of the story for real r contained in the work of Bourdon et al. in [3], we now know that symbols belonging to the extended Cowen–Kriete family induce cohyponormal composition operators if and only if $0 \leqslant r \leqslant 1$.

In the final section of the paper, we consider a detailed example and establish connections between extremal non-compactness and another heavily studied operator-theoretic quantity that is often compared to $\|C_\varphi\|$.

6. The quantity S_φ^* and geometric function theory

Another way to try to compute $\|C_\varphi\|$ that has drawn a fair amount of attention recently is to compare it to the quantity

$$S_\varphi^* := \sup_{w \in \mathbb{D}} \left\| C_\varphi^* \left(\frac{K_w}{\|K_w\|} \right) \right\|,$$

where K_w represents the reproducing kernel at w in H^2. While it is clear that $\|C_\varphi\| \geqslant S_\varphi^*$, this inequality may be strict. For example, if $\varphi(z) = 2/(3 - z)$, then $\|C_\varphi\| > S_\varphi^*$ ([2, Theorem 4.2] or [15]). On the other hand, if $\varphi(z) = sz + t$ for constants satisfying $|s| + |t| \leqslant 1$, then $\|C_\varphi\| = S_\varphi^*$ (see [6, Proof of Theorem 3], for the case $|s| + |t| = 1$; see [4, Proof of Theorem 3.2], for the case $|s| + |t| < 1$).

In [4], the second author and Paul Bourdon show that for φ with $\varphi(0) \neq 0$ inducing compact C_φ, $\|C_\varphi\| = S_\varphi^*$ if and only if $\varphi(z) = sz + t$. As a corollary, non-affine maps φ inducing C_φ with $\|C_\varphi\| = S_\varphi^*$ are shown to satisfy $\|C_\varphi\| = \|C_\varphi\|_e$.

As evidenced by Joel Shapiro's celebrated formula for the essential norm of a composition operator on H^2 (see [16])

$$\|C_\varphi\|_e^2 = \limsup_{|w| \to 1^-} \frac{N_\varphi(w)}{-\log|w|},$$

the compactness of C_φ is intimately related to the way in which φ approaches the boundary of the disk, both how fast and how often. Using this formula, it is not difficult to show that for univalent φ, $S_\varphi^* \geqslant \|C_\varphi\|_e$. Now, since affine maps induce C_φ satisfying either $\|C_\varphi\|_e = 0$ or $\|C_\varphi\|_e = \|C_\varphi\|$, the preceding results combine to yield the following fact.

Fact. If φ is univalent, $\varphi(0) \neq 0$, and C_φ is non-compact, then $\|C_\varphi\| = S_\varphi^*$ if and only if $\|C_\varphi\| = \|C_\varphi\|_e$.

The complete classification of extremal non-compactness of composition operators induced by linear fractional maps provided in the previous section therefore completes the classification of such composition operators satisfying $\|C_\varphi\| = S_\varphi^*$ as well. Theorem 4.4 adds to the evidence provided in [4] that the set of normalized reproducing kernels is rarely robust enough to exhibit the norm of a composition operator under the action of the adjoint C_φ^*.

In the case of linear fractional φ, there is a particularly nice geometric interpretation of the rarity of this phenomenon. Let us refer to a linear fractional map φ, with image disk $\Delta = \varphi(\mathbb{D})$ tangent at $\eta \in \partial\mathbb{D}$, as *fast* if for some (hence, all) $r \in (0, 1)$ we have

$$d\big(\partial\Delta, \varphi(\Gamma_r)\big) = d\big(\eta, \varphi(\Gamma_r)\big),$$

where $\Gamma_r = \{z: |z| = r\}$ and d is Euclidean distance.

Because the sets Γ_r are left invariant under rotations, the same reasoning found at the beginning of Section 4 allows us to focus our attention on maps φ given by

$$\varphi(z) = \frac{(1 + q + qd)z + (d - q - qd)}{z + d}.$$

In this case, the condition $d > 0$ is precisely the condition that φ be fast. Thus, for non-affine, non-automorphic φ that fix the point 1, $\|C_\varphi\| = S_\varphi^*$ if and only if $\|C_\varphi\| = \|C_\varphi\|_e$ if and only if φ is fast. This is intuitively appealing because fast linear fractional maps approach the boundary of the disk "as fast as possible" and so should be "as non-compact as possible."

We conclude with an example. Let $\varphi(z) = 2/(3 - z)$. This non-affine, non-automorphic linear fractional self-map of the disk already fixes the point 1 and so has "qd-representation"

$$\varphi(z) = \frac{0z + (-2)}{z + (-3)}.$$

As such, $d = -3$ and we know that $\|C_\varphi\| > \|C_\varphi\|_e$ and $\|C_\varphi\| > S_\varphi^*$. This is not new information; dating back to [2], this particular map φ has been extensively studied.

What is more interesting is to examine the family $\{\varphi_\theta\}_{0 \leqslant \theta \leqslant 2\pi}$ of linear fractional maps given by

$$\varphi_\theta(z) = -\big[\varphi(\lambda z) - 3/4\big]e^{i\theta} + 3/4.$$

Here the unimodular constant $\lambda = \varphi^{-1}((3 - e^{-i\theta})/4)$ is a rotational parameter inserted to guarantee that φ_θ fixes the point 1 for every θ.

We note that φ_0 is just the original map φ and that each φ_θ maps onto the same image disk of radius $1/4$ centered at $3/4$. But the *way* in which each φ_θ maps onto this disk depends on θ. For instance, when $\theta = \pi$ the map

$$\varphi_\pi(z) = \frac{(3/2)z + 5/2}{z + 3}$$

is fast. Thus, since $d = 3$ and $q = 1/8$, we have $\|C_\varphi\|^2 = \|C_\varphi\|_e^2 = (S_{\varphi_\pi}^*)^2 = 8$.

On the other hand, the map

$$\varphi_{\frac{\pi}{2}}(z) = \frac{(15 + 15i)z + (-31 + 33i)}{20z + (-36 + 48i)}$$

is not fast but is "faster" than the original map $\varphi(z) = 2/(3 - z)$. In this case, we use the hypergeometric representation $_2F_1((-2 + 6i)/15, (6 + 2i)/5; 16/15; x)$ to numerically compute

$$\|C_{\varphi_{\frac{\pi}{2}}}\|^2 \approx 3.3764 > 16/5 = \|C_{\varphi_{\frac{\pi}{2}}}\|_e^2.$$

In general,

$$\|C_{\varphi_\theta}\|_e^2 = \frac{16}{5 + 3\cos\theta}$$

and the quotient function $Q(\theta) = \|C_{\varphi_\theta}\|/\|C_{\varphi_\theta}\|_e$ tends to 1 as θ tends to π. Indeed, $\varphi_\theta = \eta \circ \alpha_\theta$ where $\eta(z) = (z + 3)/4$ and $\alpha_\theta(z) = 4\varphi_\theta(z) - 3$ is an automorphism. It is not hard to check that $\alpha_\theta(0) = e^{i\theta}/3$ and therefore $\|C_{\alpha_\theta}\|^2 = 2$. The elementary inequalities

$$1 \leqslant \frac{\|C_{\varphi_\theta}\|^2}{\|C_{\varphi_\theta}\|_e^2} \leqslant \frac{\|C_{\alpha_\theta}\|^2\|C_\eta\|^2}{\|C_{\varphi_\theta}\|_e^2} = \frac{5 + 3\cos(\theta)}{2}$$

then readily show that $\lim_{\theta \to \pi} Q(\theta) = 1 = Q(\pi)$.

We leave the reader with this final challenge, which to our knowledge has not been met.

Challenge: Write the function Q in closed form as a (probably) transcendental function of θ.

References

[1] G.E. Andrews, R. Askey, R. Roy, Special Functions, Cambridge Univ. Press, 1999.
[2] M. Appel, P.S. Bourdon, J. Thrall, Norms of composition operators on the Hardy space, Experiment. Math. 5 (1996) 111–117.
[3] P.S. Bourdon, E.E. Fry, C. Hammond, C.H. Spofford, Norms of linear-fractional composition operators, Trans. Amer. Math. Soc. 356 (2004) 2459–2480.
[4] P.S. Bourdon, D.Q. Retsek, Reproducing kernels and norms of composition operators, Acta. Sci. Math. (Szeged) 67 (2001) 555–562.
[5] C.C. Cowen, Composition operators on spaces of analytic functions: A status report, in: Operator Theory: Operator Algebras and Applications, Part 1, Durham, NH, 1988, in: Proc. Sympos. Pure Math., vol. 51, Amer. Math. Soc., Providence, RI, 1990, pp. 131–145.
[6] C.C. Cowen, Linear fractional composition operators on H^2, Integral Equations Operator Theory 11 (1988) 151–160.
[7] C.C. Cowen, T.L. Kriete, Subnormality and composition operators on H^2, J. Funct. Anal. 81 (1988) 298–319.
[8] C.C. Cowen, B.D. MacCluer, Some problems in composition operators, in: Studies on Composition Operators, Laramie, 1996, in: Contemp. Math., vol. 213, Amer. Math. Soc., Providence, RI, 1998, pp. 17–25.
[9] C.C. Cowen, B.D. MacCluer, Composition Operators on Spaces of Analytic Functions, CRC Press, Boca Raton, FL, 1995.
[10] K.W. Dennis, Co-hyponormality of composition operators on the Hardy space, Acta. Sci. Math. (Szeged) 68 (2002) 401–411.
[11] A. Erdélyi (Ed.), Higher Transcendental Functions, vol. I, McGraw–Hill, 1953.
[12] C. Hammond, On the norm of a composition operator with linear fractional symbol, Acta. Sci. Math. (Szeged) 69 (2003) 813–829.
[13] J.E. Littlewood, On inequalities in the theory of functions, Proc. London Math. Soc. 23 (1925) 481–519.
[14] E.A. Nordgren, Composition operators, Canad. J. Math. 20 (1968) 442–449.
[15] D.Q. Retsek, The kernel supremum property and norms of composition operators, thesis, Washington University, 2001.
[16] J.H. Shapiro, The essential norm of a composition operator, Ann. of Math. 125 (1987) 375–404.
[17] J.H. Shapiro, What do inner functions know about composition operators? Monatsh. Math. 130 (2000) 57–70.