GEOMETRIC MODELS AND COMPACTNESS OF COMPOSITION OPERATORS

By Joel H. Shapiro, Wayne Smith and David A. Stegenga

Abstract. This work explores some of the terrain between functional equations, geometric function theory, and operator theory. It is inspired by the fact that whenever a composition operator or one of its powers is compact on the Hardy space H^2, then its eigenfunctions cannot grow too quickly on the unit disc. The goal is to show that under certain natural (and necessary) additional conditions there is a converse: slow growth of eigenfunctions implies compactness. We interpret the slow-growth condition in terms of the geometry of the principal eigenfunction of the composition operator (the "Königs function" of the inducing map). We emphasize throughout the importance of this eigenfunction in providing a simple geometric model for the operator's inducing map.

1. Introduction

Each holomorphic self-mapping φ of the open unit disc U induces a linear *composition operator* C_φ, defined on the space of functions holomorphic on U by

$$C_\varphi f = f \circ \varphi \qquad (f \in H(U)).$$

Interest in this class of operators originates with Littlewood's Subordination Principle [Lit, 1925], which insures that each one restricts to a bounded operator on the Hardy space H^2 of holomorphic functions whose power series have square summable coefficients.

This paper contributes to the body of recent work that explores the connection between the function theoretic behavior of φ and the action of the induced composition operator on the Hilbert space H^2 (see [Cow] for a survey of results up to 1988). We are going to study the following variation on the compactness problem:

> *Is it possible to determine if a composition operator is compact, simply by looking at its eigenfunctions?*

Our investigation will lead to interesting connections between functional analysis, geometric function theory, and functional equations. Here is a survey of what we do, with undefined concepts and missing references deferred to the next section.

1991 *Mathematics Subject Classification.* Primary 46E15 Secondary 46B05, 30D05.

Key words and phrases. H^p-Spaces, compact composition operators, angular derivative, Schröder's equation, $BMOA$, hyperbolic distance.

This research was supported in part by the National Science Foundation. It was begun while the first author was a visiting faculty member at the University of Hawaii.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

1.1. Schröder's equation and Königs's function. The eigenvalue equation for composition operators

$$f \circ \varphi = \lambda f$$

is *Schröder's equation*, a functional equation that has been around since 1870 [Sch]. Its holomorphic solutions were first investigated by Königs [Ko], who showed in 1884 that if φ is univalent and fixes a point p of U, then there is a nonconstant function σ holomorphic and univalent on U that satisfies Schröder's equation with $\lambda = \varphi'(p)$. Note that if φ is the identity mapping on U, then any univalent mapping σ will work. To rule out this situation, we shall always assume that the range $\varphi(U)$ of φ is a proper subset of U. Königs showed that under this assumption σ is unique up to constant multiples. Moreover, *all* solutions of Schröder's equation are obtained by taking positive integer powers of this principal eigenfunction (or "Königs function") σ. More precisely, if f is a function holomorphic on U that satisfies Schröder's equation for φ, then for some positive integer n we must have $f = \sigma^n$ and $\lambda = \varphi'(p)^n$ (clearly all such choices of f and λ provide solutions). We also note that, since $\varphi(U)$ is a proper subset of U, $|\lambda| < 1$.

Stated in classical terms, the goal of this paper is to relate geometric properties of φ with those of its Königs function σ.

There is compelling motivation for doing this. Suppose φ is univalent and fixes a point of U. Let G be the image of U under the Königs function σ. Then by Schröder's equation $\sigma(p) = 0$, $\lambda \cdot G \subset G$, and the diagram below commutes, where M_λ is the mapping of multiplication by λ on G.

$$\begin{array}{ccc} U & \xrightarrow{\ \varphi\ } & U \\ \sigma \big\downarrow & & \big\downarrow \sigma \\ G & \xrightarrow[M_\lambda]{} & G \end{array} \tag{1.1}$$

Viewed in this manner, Königs's theorem makes a geometric assertion:

> *Every univalent, holomorphic self-mapping of the unit disc with a interior fixed point is equivalent to multiplication by a constant, acting on some simply connected plane domain that contains the origin.*

In diagram 1.1 we call the pair (M_λ, G) the *Königs model* for φ. From this point of view, one can understand φ by learning about G, and this is how we will analyze the compactness problem for composition operators.

We are not the first to use models to analyze composition operators. Carroll and Cowen [CarCow] recently used the Schröder model point of view as part of their construction of compact composition operators not in any Schatten class. More generally, it is known that for *any* univalent self-map φ of U the situation of diagram 1.1 obtains, with M_λ replaced by a more general linear-fractional self-map of U (see [Val] [Pomm4], [BaPomm], and [Cow2]). Such *linear-fractional models* have proved valuable in studying spectra [Cow3], co-subnormality [CowKr], and cyclicity [BoS] of composition operators.

We should point out that Königs did not require φ to be univalent, insisting only that its derivative not vanish at the fixed point. If φ is not univalent, then σ will not be univalent either (and a similar comment applies to the more general linear fractional models). While it is possible to place at least part of our work in the non-univalent setting, we choose to keep matters grounded firmly in geometry by insisting that φ, and therefore σ, always be univalent.

Here is a very simple instance of the kind of result that motivates our investigation. We use this notation: For n a positive integer, φ_n denotes the n-th iterate of φ (the composition of φ with itself n times). For f a complex function on U, the supremum of the moduli of its values on U is denoted by $\|f\|_\infty$.

Proposition. *Suppose φ fixes a point of U. If $\|\varphi_n\|_\infty < 1$ for some positive integer n, then σ is bounded on U.*

Proof. Without loss of generality we may assume that $\varphi(0) = 0$. Note that σ is still the Königs function of φ_n, the only difference being that λ^n replaces λ in Schröder's equation. By hypothesis, $\varphi_n(U)$ lies in a compact subset of U, and by Schröder's equation, $\sigma(U) = \lambda^{-n}\sigma(\varphi_n(U))$. Thus $\sigma(U)$ is bounded. □

Figure 1 shows that it may take several iterations for φ to shrink U into a compact subset of itself. Here $\varphi(z) = \sigma^{-1}(\sigma(z)/2)$, that is, φ is conjugate, via σ, to multiplication by $1/2$ on G. Note how the picture illustrates the point of view expressed by the commutative diagram (1.1) above: the action of φ is understood from the geometry of $G = \sigma(U)$.

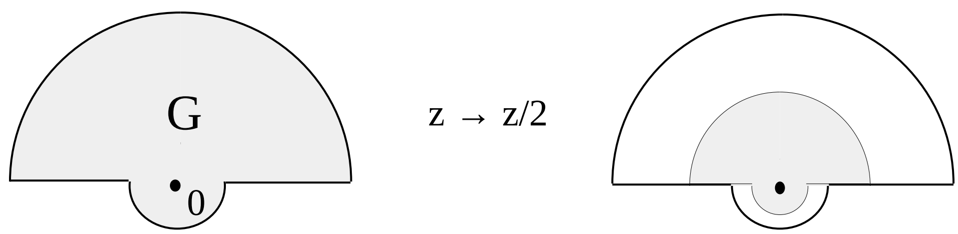

FIGURE 1. Several iterates are required to shrink G.

1.2. Schröder's equation and spectra. We will see before long that the condition $\|\varphi_n\|_\infty < 1$ guarantees that the composition operator C_φ^n is compact. Right now we observe that if this compactness is all we assume, then there results a more sophisticated smallness condition on Königs's function.

To see how this comes about, suppose that for some positive integer n, the operator C_φ^n is compact on H^2. It is well known to composition operator specialists that φ then has a fixed point p in U, and that the spectrum of C_φ consists of the points 0 and 1, along with all the positive integer powers of $\varphi'(p)$ ([C-S]; we provide more details in §2). When $\varphi'(p) \neq 0$, the Riesz theory of compact operators tells us that these non-zero spectral points have to be eigenvalues of finite multiplicity. For univalent maps φ, Königs's results complete the picture: the multiplicities are all 1, and (since $C_\varphi^n = C_{\varphi_n}$) the successive eigenfunctions are simply the positive integer powers of the univalent map σ. Thus each power of the Königs function σ belongs to H^2, so σ itself must therefore belong to each Hardy space H^p ($p < \infty$). In summary:

Folk-Theorem. *If some power of a composition operator is compact on H^2, then the inducing map has an interior fixed point, and its Königs function belongs to H^p for every $p < \infty$.*

The goal of this paper is to investigate the *converse* of this folk-theorem. In order to understand the issues involved, we need to review some of the intuition for how one uses properties of φ to decide if C_φ is compact.

1.3. The compactness problem. Recall that a Hilbert space operator is *compact* if it maps the closed unit ball of the Hilbert space into a compact set. Thus on infinite dimensional Hilbert space, where balls are not compact, such operators have to do a lot of compressing. The problem for *composition operators* is to relate this to the way in which the inducing maps compress the unit disc into itself.

A couple of elementary cases (discussed in more detail below, and in §2) set the stage:

(a) If the image of φ has compact closure in U (i.e., $\|\varphi\|_\infty < 1$), then C_φ is compact.
(b) At the other extreme, if φ has radial limits of modulus 1 on a set of positive measure on ∂U (e.g., if φ is the identity map on U), then C_φ is not compact.

Two intermediate cases lend substance to the compactness problem:

(c) If $\varphi(z) = (1 + z)/2$, so the image of the unit disc is a disc in U that is tangent to the unit circle at the point 1, then C_φ is *not* compact.
(d) On the other hand, if φ maps the unit disc into a polygon inscribed in the unit circle, then C_φ *is* compact.

1.4. First taste of the main result. The results of §1.1 and §1.2 tell us that if some power of a composition operator is compact, then the Königs function of the inducing map must be small. Here is a simple result that shows how one can go in the other direction. It is the converse for univalent maps of the Proposition of §1.1.

Proposition. *Suppose the univalent map φ has a fixed point in U, and its Königs function σ is bounded on U. Then $\|\varphi_n\|_\infty < 1$ for some positive integer n.*

Proof. We may suppose without loss of generality that $\varphi(0) = 0$. Since σ is univalent, we can call on the continuity of σ^{-1}, and the fact that $\sigma(0) = 0$, to provide a positive number δ such that $|\sigma(z)| < \delta$ implies $|z| < 1/2$. Since σ is bounded on U, there exists a positive integer n such that $\lambda^n \sigma(U) \subset \delta U$ (recall that we always have $|\lambda| < 1$). Schröder's equation (for φ_n) now yields:

$$\sigma(\varphi_n(U)) = \lambda^n \sigma(U) \subset \delta U,$$

hence our choice of δ insures that $\varphi_n(U) \subset U/2$. □

1.5. Compactness and angular derivatives. The complete characterization of compact composition operators on H^2 involves issues of value distribution theory that needn't concern us here [Sh]. For univalent inducing maps the situation simplifies considerably ([M-S, Theorem 3.10] and §2 below):

> *For φ univalent; C_φ is compact on H^2 if and only if φ has an angular derivative at no point of ∂U.*

In this paper, when we say that φ has an *angular derivative* at a point $\omega \in \partial U$, we mean that as $z \to \omega$ non-tangentially, two things happen: (i) φ has non-tangential limit η *of modulus one*, and (ii) the difference quotient

$$\frac{\eta - \varphi(z)}{\omega - z}$$

has a (finite) limit. It is not difficult to see that the second condition is equivalent to the existence of the limit of the derivative $\varphi'(z)$ as $z \to \omega$ non-tangentially [Car, pp. 298–304]. This latter condition is often taken by itself as the definition of angular derivative (cf. [Pomm1, §10.2]). Thus our definition is somewhat restrictive in requiring the existence of the angular derivative to also force the image of the unit disc out to the boundary. For example it mandates that if $\|\varphi\|_\infty < 1$, then φ has an angular derivative at *no* point of ∂U, even if φ is holomorphic across the entire unit circle. The point here is that, in the study of composition operators, the only interesting phenomena are those that occur as the image approaches the boundary.

In addition, our definition insures that the angular derivative is never zero, so the same proof that shows "analyticity implies conformality" in the interior also applies at the boundary, and shows that the existence of the angular derivative implies a form of non-tangential boundary conformality (see §2 for the precise assertions).

To appreciate the utility of this angular derivative criterion for compactness, the reader might find it instructive to use it verify that the maps $\varphi(z) = (1 \pm z)/2$ induce non-compact composition operators, as does any conformal map of the unit disc onto a subdomain whose boundary contains an arc of the unit circle, while the conformal map of the unit disc onto an inscribed polygon induces a compact operator.

The angular derivative criterion allows an entirely classical rendering of the folk-theorem of §1.2:

> *If for some iterate of a univalent self-map of the disc the angular derivative exists at no boundary point, then the original map has an interior fixed point, and its Königs function belongs to every H^p space ($p < \infty$).*

Here is an example that illustrates how this can happen, even if $\|\varphi\|_\infty = 1$. Let

$$\sigma(z) = \log \frac{1+z}{1-z},$$

so σ maps the unit disc onto the strip $\{|\Im z| < \pi/2\}$. For $0 < \alpha < 1$ let $\varphi_\alpha = \sigma^{-1} \circ M_\alpha \circ \sigma$, so φ_α maps U onto the "lens" that is symmetric about the real axis, and whose boundary makes angle $\alpha\pi/2$ radians with that axis at the points ± 1 (Figure 2). We call φ_α a "lens map."

Since non-tangential conformality fails for φ_α at the fixed points ± 1, it does not have an angular derivative at these points (they are fixed points for the map). Since these are the only places the map could hope to have an angular derivative, it follows that it induces a compact composition operator on H^2. Note that σ is in H^p for every $p < \infty$ (in fact, σ is the archetypal $BMOA$ function; see §5 below).

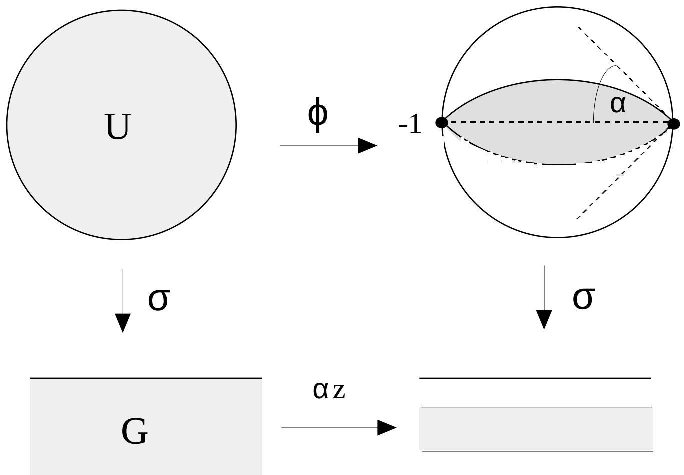

FIGURE 2. The lens map φ_α.

1.6. The converse fails. Here is a simple example that shows that, if we wish to prove a converse to the folk-theorem, then some extra hypotheses will be needed.

Let G denote the union of the unit disc and the horizontal strip $\{0 < \Im z < 2\}$. The Riemann Mapping Theorem produces a univalent map σ that takes U onto G, with $\sigma(0) = 0$ and $\sigma'(0) > 0$. We define a univalent self-map φ on U by the equation

$$\varphi(z) = \sigma^{-1}(\frac{1}{2}\sigma(z)).$$

Then σ is the Königs function of φ with $\lambda = 1/2$, and the n-th iterate of φ is obtained by replacing the multiplier $1/2$ by $(1/2)^n$ in the definition of φ. As Figure 3 shows, each of these iterates maps some interval of the unit circle onto another such interval, so by (b) of §1.3 no power of C_φ is compact on H^2.

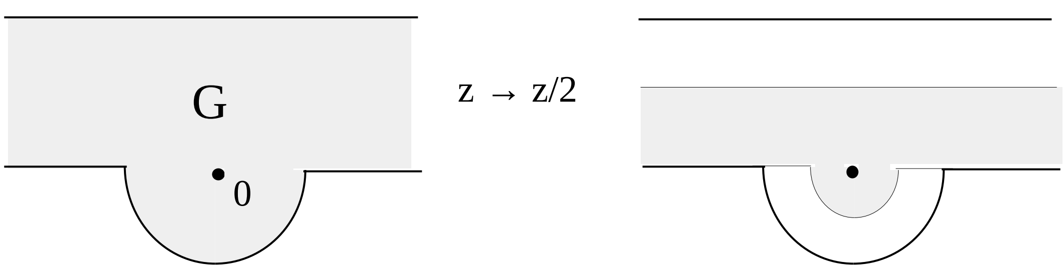

FIGURE 3. All iterates of φ induce noncompact operators.

On the other hand, $G = \sigma(U)$ is contained in a strip, so a standard subordination argument shows that σ belongs to H^p for every $p < \infty$ (and even to $BMOA$). To summarize:

> *For a univalent function φ with fixed point in U, the Königs function may belong to every H^p $(p < \infty)$, even though no power of C_φ is compact.*

1.7. Main results. In the example above, each iterate of φ has an angular derivative on a complete arc of the unit circle. However it does not have an angular derivative at any *boundary fixed point*. As the result below (which will eventually be Proposition 3.3) shows, therein lies the key.

> *Suppose φ is univalent, with a fixed point in U. If φ fixes a point of ∂U (in the sense of radial limits) at which it has an angular derivative, then $\sigma(U)$ contains a "twisted sector."*

Definition. A domain G *contains a twisted sector* if there exists an unbounded curve $\gamma \subset G$ and a positive constant ϵ such that

$$\delta_G(w) \geq \epsilon |w| \qquad w \in \gamma.$$

Here $\delta_G(w)$ denotes the distance from $w \in G$ to ∂G.

A standard subordination argument implies that whenever a simply connected domain contains an angular sector, the Riemann mapping function taking the unit disc onto that domain must fail to belong to some H^p space. We will prove that the same is true for twisted sectors, and this will provide the desired H^p estimates for the Königs function. This "twisted" subordination result, along with Proposition 3.3 above yields our main results. For their statements, we remind the reader that φ is always a univalent self-map of U that fixes the origin, $\lambda = \varphi'(0)$, σ is the *Königs map of* φ (the univalent mapping that shows up diagram 1.1, and $G = \sigma(U)$, the *Königs domain* of φ.

First Main Theorem. *Suppose φ is a univalent self-map of U with a fixed point in U, and that for some positive integer n_0 there are at most finitely many points of ∂U at which φ_{n_0} has an angular derivative. Then the following are equivalent:*

(a) *Some power of C_φ is compact on H^2;*
(b) *σ lies in H^p for every $p < \infty$;*
(c) *G does not contain a twisted sector.*

It is useful to replace the initial hypothesis on φ by a geometric condition on the image of φ. For example, it is enough to demand that the closure of $\varphi(U)$ intersect the boundary of U at only finitely many points. More generally, the boundary conformality that follows from the definition of angular derivative insures that if φ has an angular derivative at a boundary point ω then $\varphi(\omega)$ is *angularly accessible* from $\varphi(U)$, in the following sense:

Definition. A point $\zeta \in \partial U$ is *angularly accessible* from $V \subset U$ provided that given any angle $\Gamma_\alpha(\zeta)$ based at ζ, some truncation $\Gamma_\alpha(\zeta) \cap \{z \in U \mid |z| > r\}$ is contained in V for some $r < 1$. Here $\Gamma_\alpha(\zeta)$ denotes the convex hull of the point ζ and the disk $\{z \in U \mid |z| < \sin\alpha\}$.

In Figure 4, just one point is angularly accessible from $\varphi(U)$, even though the closure of that image contacts ∂U on an entire arc.

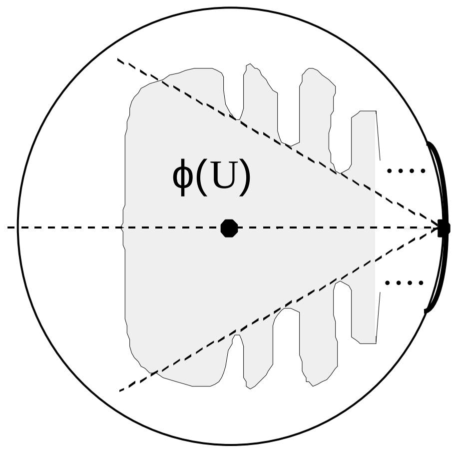

FIGURE 4. Only one point is angularly accessible from $\varphi(U)$.

Corollary 1. *Suppose that, for some positive integer n_0, at most finitely many points of ∂U are angularly accessible from $\varphi_{n_0}(U)$. Then the following are equivalent:*

(a) *Some power of C_φ is compact on H^2;*
(b) *σ lies in H^p for every $p < \infty$;*
(c) *G does not contain a twisted sector.*

In keeping with our view that univalent self-maps of the disc that fix an interior point are modeled, via Königs's theorem, by multiplication by a constant, it would be nice to be able to decide if C_φ is compact by simply looking at a picture of $G = \sigma(U)$. In this regard, if we rule out the type of behavior exhibited in §1.6, then it is enough to assume that G has *at most finitely many prime ends at* ∞ (i.e., at most finitely many hyperbolic geodesics from the origin to ∞).

Corollary 2. *Suppose that $\lambda \cdot \partial G \subset G$ and that G has only finitely many prime ends at ∞. Then the following are equivalent:*

(a) *Some power of C_φ is compact on H^2;*
(b) *σ lies in H^p for every $p < \infty$;*
(c) *G does not contain a twisted sector.*

The assumption in the First Main Theorem that there are at most finitely many points where φ_n has as angular derivative is necessary. The same is true for the corresponding finiteness assumptions in Corollary 1 and Corollary 2. An example proving this will be provided later. However, our methods do provide the same kind of result when $\varphi(U)$ meets the boundary infinitely often, provided G is "strictly starlike."

Definition. A simply connected domain G is *strictly starlike* (with respect to the origin) if $tw \in G$ for every $w \in \overline{G}$ and every $0 < t < 1$.

Now suppose as usual that φ is a univalent self-map of U with an interior fixed point, that σ is the Königs function for φ, and that $G = \sigma(U)$. Recall that bounded Königs functions are no longer an issue; we have already shown that if G is bounded then C_φ is compact.

Second Main Theorem. *Suppose G is unbounded and strictly starlike. Then there exists a positive integer n such that $\lambda^n > 0$. Let n be the least such integer; then the following are equivalent:*

(a) C_φ^n *is compact on* H^2;
(b) σ *lies in* H^p *for every* $p < \infty$;
(c) G *does not contain a sector.*

Note in particular that if $\lambda > 0$ then statement (a) becomes: C_φ *is compact.*

The First Main Theorem and its corollaries will be proved in §3, and then the example showing the finiteness assumptions are necessary will be given. §4 will be devoted to the proof of the Second Main Theorem. The Second Main Theorem turns out to be the easier one to prove; it only requires Lemma 3.2 and Proposition 3.3 of §3.

After this follows a section on the naturally related question of when the Koenigs function of φ belongs to $BMOA$ or $VMOA$. In the final section we comment on the status of the main results for some other function spaces such as the Bergman space (where it still holds) and the Dirichlet space (where it does not).

In preparation, we collect in the next section those prerequisites necessary for understanding what is to follow, along with further bibliographic references. In order to make this paper accessible to a wide audience we have tried to be reasonably complete, in some cases sketching proofs of crucial results already in the literature, but not well known outside the circle of true believers.

Acknowledgement. We wish to thank Professor Paul Bourdon of Washington and Lee University for pointing out an error in an earlier draft of this paper.

2. Background material

2.1. The Hardy spaces. For $0 < p < \infty$ the *Hardy space* H^p is the collection of holomorphic functions f on U for which

$$\|f\|_p^p \stackrel{\text{def}}{=} \sup_{0<r<1} \left\{ \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^p d\theta \right\} < \infty.$$

If $p \geq 1$ the functional $\|\cdot\|_p$ is a norm that makes H^p into a Banach space [Dur, §3.2, Cor.1]. A straightforward calculation employing the orthogonality of the monomials $\{z^n\}$ on the unit circle shows that H^2 coincides with the space of square-summable power series. More precisely, if $f(z) = \sum a_n z^n$ is holomorphic on U, then $f \in H^2$ if and only if $\sum |a_n|^2 < \infty$, in which case $\|f\|_2 = \left(\sum |a_n|^2\right)^{1/2}$.

For $0 < p < \infty$, each $f \in H^p$ obeys the growth restriction

$$|f(z)| \leq C_p \frac{\|f\|_p}{(1 - |z|)^{1/p}} \qquad (z \in U),$$

where C_p is a constant independent of f and z [Dur, §5.5, Theorem 5.9]. For $p = 2$ this condition follows readily from the power series characterization of H^2 and the Cauchy-Schwarz inequality.

2.2. Schröder's Equation. Suppose φ is a holomorphic self-map of U that fixes the origin, with $0 \neq |\varphi'(0)| < 1$ (in particular this holds if φ is univalent but not an automorphism). We are interested in solutions of Schröder's equation

$$f \circ \varphi = \lambda f, \tag{2.1}$$

where f is holomorphic on U and λ is a complex number.

Königs's Theorem. *The sequence of functions*

$$\sigma_k(z) \stackrel{\text{def}}{=} \frac{\varphi_k(z)}{\varphi'(0)^k}$$

converges uniformly on compact subsets of U to a non-constant function σ that solves (2.1) with $\lambda = \varphi'(0)$. More generally if f and λ solve (2.1), then there is a positive integer n such that $\lambda = \varphi'(0)^n$ and f is a constant multiple of σ^n.

Proof. (cf. [Ko]) The convergence of the sequence $\{\sigma_k\}$ follows from arguments involving the Schwarz Lemma. To see how this goes, let $F(z) = \varphi(z)/\varphi'(0)z$, so that

$$\sigma_k(z) = z \prod_{j=0}^{k-1} F(\varphi_j(z)),$$

where $\varphi_0(z) \equiv z$. To prove the desired convergence of $\{\sigma_k\}$ we need only prove that the series

$$\sum_{j=0}^{\infty} |1 - F(\varphi_j(z))| \tag{2.2}$$

converges uniformly on compact subsets of U. Since F is a bounded holomorphic function on U with $F(0) = 1$, the Schwarz Lemma shows that

$$|1 - F(z)| \leq A|z| \qquad (z \in U),$$

where $A = 1 + |\varphi'(0)|^{-1}$.

Now suppose $0 < r < 1$ and let $K = \{z \in U : |z| \leq r\}$. Another application of the Schwarz Lemma, this time to φ itself, shows that $\varphi(K) \subset K$, and moreover there exists a constant $m < 1$ such that for each $z \in K$ we have $|\varphi(z)| \leq m|z|$. Thus we can iterate the last inequality, obtaining for each positive integer j,

$$|\varphi_j(z)| \leq m^j|z| \qquad (z \in K).$$

From the last two displayed inequalities,

$$|1 - F((\varphi_j(z))| \leq A|\varphi_j(z)| \leq Am^j|z|$$

for each $z \in K$. Thus the series (2.2), and therefore the sequence $\{\sigma_k\}$, converges uniformly on K, as desired.

Since $\sigma_n \circ \varphi = \lambda \sigma_{n+1}$, the limit function σ obeys Schröder's equation (2.1) with $\lambda = \varphi'(0)$. Note that the fact that $\varphi'(0) \neq 1$ forces σ to be non-constant.

As for the uniqueness, suppose f and λ are any solutions of (2.1), with f holomorphic and non-constant on U. Then $\lambda \neq 1$ (else f would be constant), so upon evaluating both sides of (2.1) at the origin, we see that $f(0) = 0$. Thus there is a positive integer n and a holomorphic function g on U such that $g(0) \neq 0$ and $f(z) = z^n g(z)$ on U. This equation along with (2.1) yields for each $z \in U$:

$$\lambda z^n g(z) = \lambda f(z) = f(\varphi(z)) = \varphi(z)^n g(\varphi(z)),$$

whereupon

$$\lambda g(z) = \left(\frac{\varphi(z)}{z}\right)^n g(\varphi(z)) \qquad (z \in U,\ z \neq 0). \tag{2.3}$$

Upon letting $z \to 0$ in this equation we see that $\lambda g(0) = \varphi'(0)^n g(0)$, so because $g(0) \neq 0$ we conclude that $\lambda = \varphi'(0)^n$.

Upon substituting this into (2.3) and recalling that $f(z) = z^n g(z)$ we obtain

$$f(z) = \left(\frac{\varphi(z)}{\varphi'(0)}\right)^n g(\varphi(z)).$$

Now substitute $\varphi_j(z)$ for z in the last equation and use (2.1) one last time to obtain for each positive integer j:

$$\lambda^j f(z) = f(\varphi_j(z)) = \left(\frac{\varphi_{j+1}(z)}{\varphi'(0)}\right)^n g(\varphi_{j+1}(z)).$$

Thus for each j,

$$f(z) = \left(\frac{\varphi_{j+1}(z)}{\varphi'(0)^{j+1}}\right)^n g(\varphi_{j+1}(z)) \qquad (z \in U), .$$

Now let $j \to \infty$ and recall that $\varphi(0) = 0$. An argument involving the Schwarz Lemma shows that because of this, $\varphi_n(z) \to 0$, hence $f(z) = g(0)\sigma(z)^n$, as desired. $\square$

Note that if φ is univalent, then so is each normalized iterate σ_n, and therefore by Hurwitz's Theorem, so is σ.

2.3. The Julia-Carathéodory Theorem. Suppose as usual that φ is a holomorphic self-map of U and that $\omega \in \partial U$. The Julia-Carathéodory Theorem asserts that the following three statements are equivalent:

(a) $\liminf_{z \to \omega} \dfrac{1 - |\varphi(z)|}{1 - |z|} = \delta < \infty$.

(b) $\angle \lim_{z \to \omega} \dfrac{\eta - \varphi(z)}{\omega - z}$ exists for some $\eta \in \partial U$.

(c) $\angle \lim_{z \to \omega} \varphi'(z)$ exists, and $\angle \lim_{z \to \omega} \varphi(z) = \eta \in \partial U$.

Here $\angle\lim_{z\to\omega}$ denotes "non-tangential limit at ω." In addition, the Julia-Carathéodory Theorem establishes that $\delta > 0$ in part (a), that the points η in (b) and (c) are the same, and that both the derivative in (c) and the difference quotient in (b) have the same non-tangential limit, namely $\omega\bar{\eta}\delta$ (see [Car, §295–§303]).

Statement (b) asserts, of course, the existence of the angular derivative previously described in §1.5. Note that the fact that the angular derivative does not vanish and has the particular form noted above shows that if it exists at ω then φ maps the radius to ω to a curve in U that approaches η tangent to the radius at η. More generally, any curve in U that ends at ω and makes an angle $0 \le \alpha < \pi/2$ with the radius to ω is taken into a curve that makes the same angle with the radius to η. This is the "boundary conformality" mentioned in §1.5.

2.4. The compactness problem – univalent case.

Here is the intuition behind the "univalent compactness theorem" of §1.5. A straightforward calculation with power series shows that for $f(z) = \sum a_n z^n$ holomorphic on U,

$$\sum |a_n|^2 \approx |f(0)|^2 + \int_U |f'(z)|^2(1 - |z|^2)\, dA(z) \tag{2.4}$$

where dA represents Lebesgue area measure on U, and "$\approx$" means that the left hand side is bounded above and below by constant multiples of the right hand side, where the constants are independent of f. Thus f is in H^2 if and only if the integral on the right side of (2.4) converges, and if we ignore the inessential term $f(0)$, the norm of f is essentially comparable to the square root of that integral. In particular if $f \in H^2$ then the change of variable formula shows that for φ a univalent self-map of U,

$$\|f \circ \varphi\|_2^2 \le \text{const.} \left(|f(\varphi(0))|^2 + \int_{\varphi(U)} |f'(w)|^2 \Omega_\varphi(w)(1 - |w|^2)\, dA(w) \right),$$

where

$$\Omega_\varphi(w) \overset{\text{def}}{=} \frac{1 - |z|^2}{1 - |\varphi(z)|^2} \qquad (w = \varphi(z)).$$

Suppose for simplicity (and without loss of generality) that $\varphi(0) = 0$. Then the Schwarz Lemma insures that $\Omega_\varphi(w) \le 1$ for every $\omega \in \varphi(U)$, and this establishes the boundedness, at least for univalent φ, of the operator C_φ. It is often the case in operator theory that if a "big-oh" condition implies boundedness for a class of operators, then the corresponding "little-oh" condition will imply compactness. This is exactly what happens here: it turns out that C_φ is compact if and only if $\Omega_\varphi(w) \to 0$ as $w \to \partial\varphi(U)$, i.e., if and only if

$$\lim_{|z|\to 1-} \frac{1 - |z|^2}{1 - |\varphi(z)|^2} = 0,$$

and by the Julia-Carathéodory Theorem this is the same as asserting that φ has an angular derivative at no point of ∂U. (For more general compactness results, see [M-S], [Sh].)

2.5. The Folk-Theorem. We discuss in more detail the fact that:

> *If C_φ is compact on H^2 then φ has a fixed point in U and Königs's function for φ lies in H^p for every finite p.*

The existence of the fixed point comes from the Denjoy-Wolff Theorem which asserts that if φ has no fixed point in U then there is a point $\omega \in \partial U$ such that $\varphi_n \to \omega$ uniformly on compact subsets of U. Furthermore this "Denjoy-Wolff point" ω has the remarkable property that the angular derivative of φ exists there! Thus if φ has no fixed point in U, then C_φ cannot be compact.

Now suppose C_φ is a compact operator. Because conformal automorphisms of U induce isomorphic composition operators on H^2, we may suppose without loss of generality that the interior fixed point of φ is at the origin: $\varphi(0) = 0$. Let $\lambda = \varphi'(0)$, and suppose for simplicity that φ is univalent, so $\lambda \neq 0$. Then the matrix of C_φ with respect to the orthonormal monomial basis $\{z^n\}_0^\infty$ is lower triangular, with the sequence $\{\lambda^n\}$ on the diagonal. Thus the adjoint operator C_φ^* is upper triangular with $\{\bar{\lambda}^n\}$ as its diagonal, hence each of these diagonal entries is an *eigenvalue* of C_φ^* (the upper triangularity is crucial here). Thus the numbers λ^n belong to the spectrum of the original compact operator C_φ, and since they are non-zero, the Riesz theory of compact operators guarantees that each is an eigenvalue of C_φ. In other words, for each positive integer n there is a non-zero vector $f \in H^2$ such that

$$f \circ \varphi = \lambda^n f.$$

But Königs's Theorem guarantees that f is a constant multiple of σ^n, where σ is the Königs function produced in §2.2. Thus if C_φ is compact, then every power of σ belongs to H^2, so σ itself lies in H^p for every finite p.

2.6. The Hyperbolic metric. Let ρ_U denote the hyperbolic metric on U, defined by (see [A, p. 2])

$$\begin{aligned} \rho_U(z_1, z_2) &= \inf \left\{ \int_\gamma \frac{2|dz|}{1-|z|^2} : \gamma \text{ is an arc in } U \text{ from } z_1 \text{ to } z_2 \right\} \\ &= \log \frac{|1-\overline{z_1}z_2| + |z_1 - z_2|}{|1-\overline{z_1}z_2| - |z_1 - z_2|}. \end{aligned}$$

This distance is invariant under conformal self-mappings of the disk, and therefore transfers to a natural conformally invariant metric on any simply connected proper subset $G \subset \mathbb{C}$. If $\sigma : U \to G$ is any conformal map, this *hyperbolic distance* on G is given by $\rho_G(w_1, w_2) = \rho_U(z_1, z_2)$, where $w_i = \sigma(z_i)$ for $i = 1, 2$. The function

$$h_G(w) = \frac{2}{|\sigma'(z)|(1-|z|^2)}, \qquad \text{where } w = \sigma(z),$$

satisfies

$$h_G(w)|dw| = \frac{2|dz|}{1-|z|^2}, \qquad w = \sigma(z).$$

Thus, ρ_G can be computed by integrating $h_G(w)$ over arcs in G. However, $h_G(w)$ is not explicitly computable in terms of G alone. A useful substitute is the *quasi-hyperbolic metric* on G, introduced by Gehring and Palka [GP]. For a domain $G \subsetneqq \mathbb{R}^n$ and $x \in G$, let $\delta_G(x)$ denote the Euclidean distance from x to the boundary of G. The quasi-hyperbolic distance from x_1 to x_2 in G is defined to be

$$k_G(x_1, x_2) = \inf \left\{ \int_\gamma \frac{ds}{\delta_G(x)} : \gamma \text{ is an arc in } G \text{ from } x_1 \text{ to } x_2 \right\}.$$

Here ds denotes integration with respect to arclength.

The quasi-hyperbolic metric is closely related to the hyperbolic metric. Indeed, if G is a simply connected domain in the complex plane, then

$$\frac{1}{2} \le h_G(w)\delta_G(w) \le 2$$

[Pomm1, p. 22]. It follows that

$$\frac{1}{2}\rho_G \le k_G \le 2\rho_G\,.$$

Due to the geometric nature of its definition, k_G is thus very useful in obtaining estimates for the hyperbolic metric.

3. Proof of First Main Theorem and Corollaries.

The following variation on the theme of the First Main Theorem is the crucial step in its proof. For notational convenience, if $\omega \in \partial U$, we write

$$\varphi(\omega) = \lim_{r \to 1-} \varphi(r\omega)$$

whenever the radial limit on the right exists. If $\varphi(\omega) = \omega$, we say ω is a *boundary fixed point* of φ.

3.1. Theorem. *Suppose φ is univalent, not an automorphism of U, and has a fixed point in U. Suppose further that φ has a finite angular derivative at some* boundary fixed point *ω. Then for some $p < \infty$ the Königs function of φ does not belong to H^p.*

The proof requires some preliminaries, the first of which is a Euclidean lower bound on the quasi-hyperbolic metric. Recall that if G is a domain, and $z \in G$, then $\delta_G(z)$ denotes the distance from z to ∂G. This simple, but useful estimate occurs in the work of Gehring and Palka [Ge-Pal, Lemma 2.1]. In order to keep our exposition complete, we provide a proof.

3.2. Lemma. *If a and b are points of a domain G, then*

$$k_G(a,b) \geq \log\left(1 + \frac{|a-b|}{\min(\delta_G(a), \delta_G(b))}\right).$$

Proof. Let Γ be any rectifiable curve in G from a to b,where a is the endpoint closest to the boundary of G. For $\zeta \in \Gamma$ let s denote the arc length on Γ from a to ζ and $|\Gamma|$ the arc length of Γ. The triangle inequality shows that

$$\delta_G(\zeta) \leq \delta_G(a) + |a-\zeta| \leq \delta_G(a) + s,$$

so

$$\begin{aligned} \int_\Gamma \frac{|d\zeta|}{\delta_G(\zeta)} &\geq \int_0^{|\Gamma|} \frac{ds}{\delta_G(a)+s} \\ &= \log\left(1 + \frac{|\Gamma|}{\delta_G(a)}\right) \\ &\geq \log\left(1 + \frac{|a-b|}{\delta_G(a)}\right). \end{aligned}$$

The result follows upon taking the infimum of the left hand side over all such curves Γ. □

This lemma is the key to the following geometric version of Theorem 3.1.

3.3. Proposition. *Under the hypotheses of Theorem 3.1, $G = \sigma(U)$ contains a twisted sector about the image of the radius to the boundary fixed point of φ. That is, $\lim_{r\to 1-} \sigma(r\omega) = \infty$, and there exists a positive constant ϵ such that*

$$\delta_G(\sigma(r\omega)) \geq \epsilon|\sigma(r\omega)|$$

for every $0 \leq r < 1$.

Remark. To better understand this result, suppose $\omega = 1$, $\varphi(0) = 0$, and φ is real on the real axis. Then $\varphi'(0)$ is real, so by Schröder's equation σ is real on the real axis. In this case Proposition 3.3 asserts that $\sigma(U)$ contains an *angular sector* symmetric about the real axis (with vertex at the origin). Write the angular opening of the sector as (π/p) for some $1 < p < \infty$. Consider the mapping

$$F(z) = \left(\frac{1+z}{1-z}\right)^{1/p}$$

that takes U conformally onto the sector. Then $\psi = \sigma^{-1} \circ F$ is a holomorphic self-map of U, and $F = \sigma \circ \psi$. Since the composition operator C_ψ maps each Hardy space into itself, F is in every Hardy space to which σ belongs. Since F is not in H^p, neither is σ.

In order to use Proposition 3.3 to finish the proof of Theorem 3.1, we will have to find a "twisted" substitute for the above argument.

Proof of Proposition 3.3. We may suppose without loss of generality that the interior fixed point of φ is at the origin, and that the boundary fixed point at which the angular derivative exists is the point 1. Thus

$$\lim_{r\to 1-} \frac{1-\varphi(r)}{1-r} = d \in [1,\infty).$$

(The positivity of d follows from the Julia-Carathéodory Theorem, while $d \geq 1$ is an immediate consequence of the Schwarz Lemma. In fact, the proof of the Denjoy-Wolff Theorem that shows $d > 1$, but we do not require this fact.) Thus

$$|\varphi(r) - r| \leq |1-\varphi(r)| + |1-r| \leq 3d(1-r),$$

for r sufficiently close to 1, say $r_0 < r < 1$. We also have, by the Schwarz Lemma, that $|\varphi(r)| \leq r$, and so $\delta_U(z) \geq 1-r$ for each z on the line segment from r to $\varphi(r)$. Hence, estimating the quasi-hyperbolic distance by integrating along this line segment, we get

$$\rho_U(r,\varphi(r)) \leq 2k_U(r,\varphi(r)) \leq 2\cdot\frac{3d(1-r)}{1-r} = 6d,$$

provided that $r_0 < r < 1$.

By Schröder's equation we can transfer this estimate to $G = \sigma(U)$, with $\sigma(r)$ playing the role of r, and $\lambda\sigma(r)$ cast as $\varphi(r)$:

$$\limsup_{r\to 1-} \rho_G(\sigma(r),\lambda\sigma(r)) \leq 6d\ . \tag{3.1}$$

Let us show that $\sigma(r)\to\infty$. Suppose otherwise. Then there is a sequence $0 < r_n \nearrow 1$ and a (finite) point w_0 such that $\sigma(r_n)\to w_0$. Necessarily $w_0 \in \overline{G}$. By the univalence of σ, the point w_0 cannot belong to G, so it must lie on the boundary. But according to (3.1) above, the hyperbolic distance from $\sigma(r_n)$ to $\lambda\sigma(r_n)$ has to stay bounded, and this forces $\lambda\sigma(r_n)$ to converge to w_0 also. But $\lambda\sigma(r_n)$ converges to λw_0, so $\lambda w_0 = w_0$, hence $\lambda = 1$ ($w_0 \neq 0$, since $0 = \sigma(0)$ is already in the interior of G). But this forces φ to be the identity map, contradicting our hypothesis that it is not an automorphism. Thus $\sigma(r)\to\infty$.

Now (3.1) and Lemma 3.2 show that

$$\limsup_{r\to 1-} \log\frac{|\sigma(r)-\lambda\sigma(r)|}{\delta_G(\sigma(r))} \leq 12d,$$

and hence

$$\limsup_{r\to 1-} \frac{|1-\lambda||\sigma(r)|}{\delta_G(\sigma(r))} \leq \exp(12d).$$

Since $\lambda = \varphi'(0) \neq 1$ we can rewrite this last inequality as

$$\limsup_{r\to 1-} \frac{|\sigma(r)|}{\delta_G(\sigma(r))} \leq \frac{\exp(12d)}{|1-\lambda|},$$

and this, in turn, yields the desired conclusion. □

We finish the proof of Theorem 3.1 by translating the geometric conclusion of Proposition 3.3 into a growth estimate on σ. This requires an elementary covering lemma.

3.4. Covering Lemma. *Let $\{\Delta_j\}$ be a countable collection of open disks in $\mathbb{C}$, with z_j the center of Δ_j. Suppose that $z_i \notin \Delta_j$ if $i \neq j$. Then $\sum \chi_{\Delta_j}(z) \leq 5$, for all $z \in \mathbb{C}$.*

Proof. It suffices to establish the result for $z = 0$. Suppose that $0 \in \Delta_i \cap \Delta_j$, where $i \neq j$. Then $|z_j| < |z_i - z_j|$, since $0 \in \Delta_j$ but $z_i \notin \Delta_j$. Similarly, $|z_i| < |z_i - z_j|$. Thus, by elementary trigonometry, the angle at 0 determined by z_i and z_j is (strictly) greater than $\pi/3$. Now suppose that 0 is in N of the disks $\{\Delta_j\}$. The centers of these disks determine N angles at 0, each greater than $\pi/3$ and with sum 2π. Thus $N \leq 5$, and the result follows. □

Completion of Proof of Theorem 3.1. Let γ denote the image under σ of the ray from 0 to ω. We have in hand Proposition 3.3, so γ is an unbounded curve in G that begins at the origin, and serves as the "axis" of a twisted sector that lies in G, i.e.,

$$\delta_G(w) \geq \varepsilon |w| \qquad (w \in \gamma). \tag{3.3}$$

Our intuition is that by the univalence of σ the sector cannot overlap itself, so γ cannot wrap around too tightly, and this will force the Euclidean distance from points on γ to the origin to increase rapidly as their hyperbolic distance from the origin increases. Our goal is to estimate this degree of rapidity, and then turn the result into a growth condition on σ.

Let $G_0 = G \cap U$, and for each positive integer n, set

$$G_n = G \cap \{2^{n-1} < |w| < 2^n\} \qquad \text{and} \qquad \gamma_n = \gamma \cap G_n.$$

We focus for a while on a particular segment γ_n. Observe that G_n, being a bounded open set, can be covered by a countable collection of discs

$$\Delta_j = D(\zeta_j, \frac{\delta_G(\zeta_j)}{2}),$$

where $\zeta_j \notin \Delta_i$ unless $i = j$.

To see that this is possible, first observe that we may assume that $\sigma'(0) = 1$. With this normalization of σ and the Schwarz Lemma it follows that $\delta_G(0) \leq 1$. Choose $\zeta_1 \in G_n$ so that $\delta_G(\zeta_1)$ is maximal. Having chosen $\zeta_1, \ldots \zeta_j$, choose $\zeta_{j+1} \in G_n \backslash (\Delta_1 \cup \cdots \cup \Delta_j)$ so that $\delta_G(\zeta_{j+1})$ is maximal. Clearly $\zeta_j \notin \Delta_i$ whenever $i \neq j$. We maintain that the disks $\{\Delta_j\}$ cover G_n. Indeed, the sequence $\delta_j = \delta_G(\zeta_j)$ is monotone decreasing, and since G_n is bounded, we see by the Covering Lemma that its limit is zero. By the construction, every point of $G_n \backslash (\Delta_1 \cup \cdots \cup \Delta_j)$ lies within δ_{j+1} of ∂G. Thus every point of G eventually ends up in one of the discs Δ_j.

We are not interested in all the discs Δ_j, only the ones that intersect γ_n. Let A_n denote the set of indices j for which this happens. The key to our proof is an estimate of $\#A_n$, the number of indices in A_n. This proceeds in several steps.

Observe that the condition $\delta_G(0) \leq 1$ implies that $\delta_j \leq 2^n + 1$ and hence that each disk Δ_j lies in $2^{n+2}U$. We use this fact and the Covering Lemma again to obtain these estimates:

$$\begin{aligned}\pi \sum_{j\in A_n}\left(\frac{1}{2}\delta_G(\zeta_j)\right)^2 &= \sum_{j\in A_n} \text{Area } (\Delta_j)\\ &\le 5\cdot \text{Area}\left(\bigcup_{j\in A_n}\Delta_j\right)\\ &\le 5\cdot\text{Area } (2^{n+2}U)\\ &= 5\pi\cdot 2^{2n+4}.\end{aligned}$$

Next we need to estimate $\delta_G(\zeta_j)$ for each $j \in A_n$. For such an index j we choose a point $\zeta \in \gamma_n \cap \Delta_j$, and $\eta \in \partial G$ such that $\delta_G(\zeta_j) = |\zeta_j - \eta|$. Then

$$\delta_G(\zeta_j) \ge |\zeta - \eta| - |\zeta_j - \zeta| \ge \delta_G(\zeta) - \frac{1}{2}\delta_G(\zeta_j)$$

hence

$$\delta_G(\zeta_j) \ge \frac{2}{3}\delta_G(\zeta) \ge \frac{2}{3}\varepsilon|\zeta|$$

where the last inequality comes from the twisted sector condition (3). Since ζ belongs to γ_n its modulus is $> 2^{n-1}$ so

$$\delta_G(\zeta_j) \ge \frac{2^n}{3}\varepsilon$$

for each $j \in A_n$. Along with the inequality we obtained by comparing areas, this shows that $(\#A_n)(2^{n-1}\varepsilon/3)^2 \le 5\cdot 2^{2n+4}$, and this in turn simplifies to

$$\#A_n \le \frac{2880}{\varepsilon^2}. \tag{3.4}$$

Finally, recall that each disc Δ_j has radius equal to half the distance from its center to the boundary of G. It follows easily that the quasi-hyperbolic distance between any pair of points in Δ_j is no more than 2, so the hyperbolic distance is no more than 4.

Let w_n be any point on γ with $|w_n| = 2^n$. Then w_n and w_{n-1} lie in a chain of pairwise intersecting closed discs $\bar{\Delta}_j$ with $j \in A_n$, so by (3.4) above,

$$\rho_G(w_{n-1}, w_n) \le 4(\#A_n) \le \frac{11520}{\varepsilon^2}.$$

By the triangle inequality (writing $w_0 = 0$),

$$\begin{aligned}\rho_G(0, w_n) &\le \sum_{k=1}^{n}\rho_G(w_{k-1}, w_k)\\ &\le \frac{11520}{\varepsilon^2}n\\ &= \alpha\log|w_n|,\end{aligned}$$

where $\alpha = (11520 \log 2)/\varepsilon^2$.

We finish the proof by noting that $w_n = \sigma(r_n\omega)$, where $r_n \to 1-$. By the last estimate and the conformal invariance of the hyperbolic metric,

$$\log\left(\frac{1+r_n}{1-r_n}\right) = \rho_U(0, r_n\omega) = \rho_G(0, w_n) \leq \alpha \log|w_n| = \alpha \log|\sigma(r_n\omega)|.$$

Therefore

$$|\sigma(r_n\omega)| \geq \left(\frac{1+r_n}{1-r_n}\right)^{1/\alpha} > \left(\frac{1}{1-r_n}\right)^{1/\alpha}$$

for each n. Since $r_n \to 1$, the growth estimate of §2.1 shows that $\sigma \notin H^p$ for any $p > \alpha$. □

Remarks. In view of the discussion following the statement of Proposition 3.3, one might suppose that $\sigma \notin H^p$ for any $p > C/\epsilon$, for some constant C, rather than $p > C/\epsilon^2$ as demonstrated in the proof above. This would indeed have been the case had our definition of a twisted sector (see §1.7) been based on the arclength of the portion of γ to w, rather than on $|w|$. It is easy to see that if γ satisfies the definition in §1.7 and γ spirals out to infinity, it is possible that the arclength of γ to w is comparable to $|w|/\epsilon$. Thus $p > C/\epsilon^2$ is best possible.

We also note that the proof above could as well have been based on a decomposition of G into Whitney cubes. The idea is that the minimum number of cubes in a chain of Whitney cubes containing two points of G is comparable to the quasi-hyperbolic distance between the points, provided that the distance is at least one (see [Sm-Ste, Lemma 9]).

Proof of the First Main Theorem, completed. We now prove the equivalence of (a), (b), and (c) in the statement of the First Main Theorem. That (a) implies (b) is just the statement of the Folk Theorem, while the the proof that (b) implies (c) is contained in the demonstration that Theorem 3.1 follows from Proposition 3.3. Thus it remains to show that (c) implies (a).

Recall that φ is univalent with a fixed point which we may assume is at the origin, and the hypothesis is that there is a finite integer n_0 such that there are at most finitely many points on ∂U at which φ_{n_0} has an angular derivative. We wish to show that if G does not contain a twisted sector, then the operator C_φ^n (which, we recall is just C_{φ_n}) is compact for some positive integer n. For this we assume that C_{φ_n} is not compact for any n, and derive the fact that G contains a twisted sector.

Now by the Compactness Theorem for composition operators, our assumption on φ is this: *For every positive integer n, the map φ_n has an angular derivative at some point of ∂U.* We are going to show that consequently some iterate of φ has a boundary fixed point at which it has an angular derivative, whereupon an appeal to Proposition 3.3, will complete the proof of the First Main Theorem.

To start things off, for each positive integer n let E_n denote the set of points on ∂U at which φ_n has an angular derivative. Our hypotheses are that each of these sets is non-empty and that E_{n_0} is finite. We will use various aspects of the Julia-Carathéodory theorem to prove that the sets $\{E_n\}$ have the following important properties:

(i) $E_1 \supset E_2 \supset \cdots$.

(ii) $\varphi(E_{n+1}) \subset E_n$ for each n.
(iii) φ is 1-1 on each E_n.

Before proving these properties, let's see how they yield the desired result. Property (i) asserts that $\{E_n\}$ is a decreasing sequence of non-empty sets, and, since E_{n_0} is finite, it must stabilize:

$$E_N = E_{N+1} = \cdots .$$

for some positive integer N. Thus the set $E = E_N$ has these special properties: It is non-empty, every iterate of φ has an angular derivative at each of its points, and (by (ii) and (iii)) φ acts as a permutation on E. Suppose η is any point in E. Then there exist positive integers n and k such that

$$\varphi_n(\eta) = \varphi_{n+k}(\eta) = \varphi_k(\varphi_n(\eta)).$$

Thus $\omega = \varphi_n(\eta)$ is a boundary fixed point of φ_k at which φ_k has a finite angular derivative; just what we wanted to prove.

It remains only to prove properties (i) - (iii) of the sets $\{E_n\}$.

Proof of (i). Statement (i) says that if φ_{n+1} has an angular derivative at a point $\omega \in \partial U$ then so does φ_n. According to the Julia-Carathéodory theorem, our assumption is equivalent to boundedness of the quotient $\frac{1-|\varphi_{n+1}(r\omega)|}{1-r}$ for $0 \le r < 1$. But

$$\frac{1-|\varphi_n(r\omega)|}{1-r} = \frac{1-|\varphi_n(r\omega)|}{1-|\varphi_{n+1}(r\omega)|} \cdot \frac{1-|\varphi_{n+1}(r\omega)|}{1-r} \le \frac{1-|\varphi_{n+1}(r\omega)|}{1-r},$$

where the inequality follows from the Schwarz Lemma. Thus $\frac{1-|\varphi_n(r\omega)|}{1-r}$ is also bounded, so by the Julia-Carathéodory theorem φ_n has an angular derivative at ω, as desired.

Proof of (ii). This is the statement that if φ_{n+1} has an angular derivative at ω, then φ_n has one at $\varphi(\omega)$. The existence of the radial limit $\varphi(w) \in \partial U$ follows, of course, from the fact that $E_n \subset E_1$, that is, φ has an angular derivative wherever φ_n does.

Let $\eta = \varphi(w)$. By the Julia-Carathéodory theorem, it suffices to show that

$$\liminf_{z\to\eta} \frac{1-|\varphi_n(z)|}{1-|z|} < \infty, \tag{3.5}$$

and by that same theorem, our hypothesis is equivalent to the boundedness for $0 \le r < 1$ of the quotient

$$\frac{1-|\varphi_{n+1}(r\omega)|}{1-r} = \frac{1-|\varphi_n(\varphi(r\omega))|}{1-|\varphi(r\omega)|} \cdot \frac{1-|\varphi(r\omega)|}{1-r}.$$

Now the last fraction on the right is > 1 by the Schwarz Lemma, so resulting inequality says that the quotient $\frac{1-|\varphi_n(z)|}{1-|z|}$ is bounded on the curve $z = \varphi(r\omega), 0 \le r < 1$. Since $\varphi(r\omega) \to \eta$, this curve ends at η, so (3.5) is proved.

Proof of (iii). It is enough to prove that φ is 1-1 on E_1. Here we use the conformality of φ at each point where the angular derivative exists. Fix $\omega \in E_1$ and let T be a triangle with one vertex at ω and the other two placed in U so that T is symmetric about the radius that ends at ω. Let α be the angle made by this radius with the sides of T that terminate at ω. Then $\varphi(T)$ is a curvilinear triangle in U with a vertex at $\varphi(w)$, and the corresponding sides make an angle α with the radius to $\varphi(w)$. Now suppose ω' is a different point of E_1, and T' another such triangle for ω', that does not intersect T. By the comments above, if $\varphi(\omega) = \varphi(w')$, then the images of these two triangles must intersect, contradicting the univalence of φ. Thus $\varphi(\omega) \neq \varphi(w')$, hence φ is 1-1 on E_1. □

Proof of Corollary 1. Recall that the assumption in Corollary 1 is that for some integer n_0, there are at most finitely many points on ∂U that are angularly accessible from $\varphi_{n_0}(U)$. Since univalent functions are conformal at each point of ∂U at which they have an angular derivative, it is clear that φ_{n_0} can have an angular derivative at no more than finitely many points on ∂U. Thus Corollary 1 is an immediate consequence of the First Main Theorem. □

For the proof of Corollary 2 we need a sufficient condition for the σ-image of a radius to tend to ∞. We encountered a similar situation in the proof of Proposition 3.3, where the angular derivative was involved. In the result below it is the geometry of ∂G that plays the crucial role. As usual, φ is univalent and fixes the origin, σ is its Königs function, and $G = \sigma(U)$.

3.5. Lemma. *Suppose* $\lambda{\cdot}\partial G \subset G$. *If* $\zeta \in \partial U$ *and* $\lim_{r\to 1-} |\varphi(r\zeta)| = 1$, *then* $\lim_{r\to 1-} \sigma(r\zeta) = \infty$.

Proof. Suppose $\sigma(r\zeta) \not\to \infty$. Then there is a sequence $r_n \to 1-$ and a point $w_0 \in \mathbb{C}$ such that $\sigma(r_n\zeta) \to w_0$. Necessarily w_0 lies in the closure of G, but by the univalence of σ and the fact that $r_n \to 1$, it cannot lie in G. Thus $w_0 \in \partial G$. By Schröder's equation,

$$\sigma(\varphi(r_n\zeta)) = \lambda\sigma(r_n\zeta) \to \lambda w_0,$$

and $\lambda w_0 \in G$ by our hypothesis on G. Thus

$$\varphi(r_n\zeta) \to \sigma^{-1}(\lambda w_0) \in U,$$

so $|\varphi(r\zeta)| \not\to 1$. □

Proof of Corollary 2. Recall that in Corollary 2 we are assuming $\lambda \cdot \partial G \subset G$ and G has only finitely many prime ends at ∞. Once again, we only need to show that the hypotheses of the First Main Theorem are satisfied. Consider a point $\zeta \in \partial U$ that corresponds to a prime end of G *not* at infinity; that is, suppose

$$\liminf_{r\to 1-} |\sigma(r\zeta)| < \infty.$$

By Lemma 3.5, φ can not have a radial limit of modulus 1 at ζ. In particular each point in the set $E \subset \partial U$ where φ has an angular derivative must correspond to a prime end of G at infinity, and so E is a finite set. Thus the First Main Theorem can be applied, and the proof is complete. □

We now present an example showing the necessity of the finiteness assumptions in The First Main Theorem and its corollaries.

Example. Let

$$G = \{(x, y) : |y| < 1 \text{ and } x > -1\} \setminus \bigcup_{k=1}^{\infty} \{(x, y) : |y| = 2^{-k},\ x \geq 2^k\},$$

and let σ be the conformal map $\sigma : U \to G$ with $\sigma(0) = 0$ and $\sigma'(0) > 0$. If we set $\lambda = 1/2$, then $\lambda \cdot G \subset G$, and so Schröder's equation holds, with $\varphi(z) = \sigma^{-1}(\sigma(z)/2)$.

Clearly G does not contain a twisted sector, since δ_G is bounded by 1. Also σ lies in H^p for every $p < \infty$, since σ has bounded imaginary part. We now fix a positive integer n and show that $C_\varphi^n = C_{\varphi_n}$ is *not* compact on H^2. In what follows, it is convenient to think of G as a "cubist jellyfish" whose body is the open square of edge length two centered at the origin, and whose tentacles are the rectangular strips

$$T_k = \{(x, y) : x \geq 2^{k+3} \text{ and } 2^{-k-1} < y < 2^{-k}\} \qquad (k = 0, 1, 2, \dots),$$

and their reflections in the real axis.

Denote the upper and lower horizontal boundary curves of T_k by I_k^+ and I_k^-. These curves are analytic arcs that are mapped by $z/2$ into ∂G. It follows from the Schwarz reflection principle and some elementary conformal mapping that there are open arcs J_k^+ and J_k^- on ∂U, with *common boundary point* ω_k, satisfying $\sigma(J_k^\pm) = I_k^\pm$. It follows from Schröder's equation that φ maps an open arc in ∂U centered at ω_k onto an open arc in ∂U containing ω_{k+1}. By the reflection principle φ extends to be analytic in a neighborhood of each ω_k. Thus φ has an angular derivative at each ω_k and, by the univalent compactness theorem in §1.5, C_φ is not compact.

Similar considerations show that C_φ^n is not compact for any $n \geq 1$. Note that $\varphi(U)$ is a Jordan subregion of the unit disc minus a countable collection of "ingrown hairs" whose "roots" lie on the unit circle, and converge to the point 1. See Figure 5 below.

There is no contradiction here with either the First Main Theorem or with Corollary 1 because φ_n has an angular derivative at each of the infinitely many interior points of each arc J_k, and each interior point of $\varphi_n(J_k)$ is angularly accessible from $\varphi_n(U)$. Thus the finiteness assumptions of these results are not satisfied, and indeed these assumptions are necessary.

The above example does not apply to Corollary 2, since $\lambda \cdot \partial G$ is not contained in G. However, G can be modified to produce an example showing the necessity of the finiteness assumption here as well.

Refined example. We begin with the original example, but widen out the omitted horizontal lines to rectangular channels

$$C_k = \{(x, y) : x \geq 2^k \text{ and } \frac{1}{2^k} \leq y \leq \frac{3}{2^{k+1}}\} \qquad (k = 1, 2, \dots),$$

and their reflections in the real axis. This example has the same properties as the one above, except that one obtains $\varphi(U)$ by removing from the unit disc, not a sequence of hairs, but instead a sequence of small notches with bases on the unit circle.

The "modified jellyfish" has tentacles

$$T_k = \{(x, y) : x \geq 2^{k+3} \text{ and } \frac{3}{2^{k+2}} < y < \frac{1}{2^k}\} \qquad (k = 0, 1, 2, \dots),$$

and their reflections in the real axis.

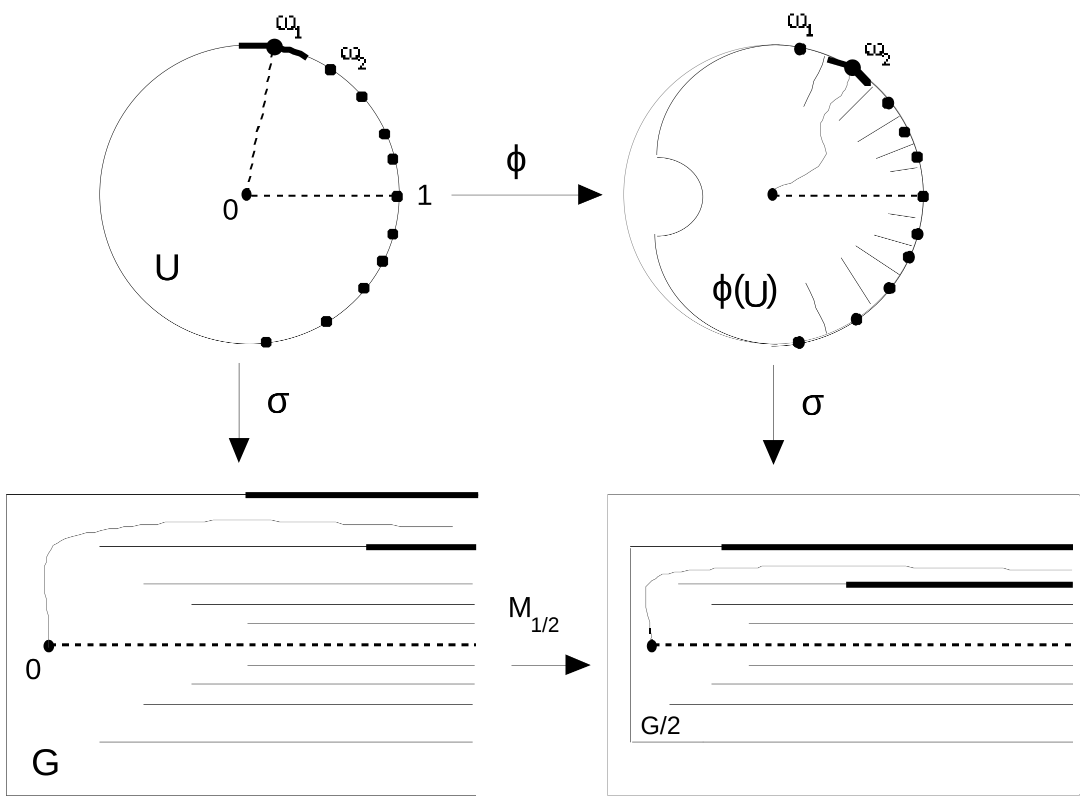

FIGURE 5. The original jellyfish; heavy curves go to heavy ones. Two radii and their images are shown dotted.

We describe a modification of the upper half of G, with the understanding that the same process is simultaneously taking place on the lower half.

As in the previous example, each tentacle corresponds to a prime end of G and the corresponding point $\omega_k \in \partial U$ has the property that the σ-image of any curve that ends at ω_k is a curve in G that tends to infinity through T_k. Moreover, φ maps an open arc in ∂U centered at ω_k onto an open arc of ∂U containing ω_{k+1} Again, the reflection principle yields that φ has a finite angular derivative at each ω_k. Let $d_k = |\varphi'(\omega_k)|$.

We now modify G so that $\frac{1}{2} \cdot \partial G \subset G$, yet the finiteness of these angular derivatives is preserved. The key is that small changes in compact parts of ∂G induce only small changes in a particular boundary point ω_k and the corresponding angular derivative $\varphi'(\omega_k)$ (we leave it as an exercise for the reader to show this).

For each integer $j \geq 1$, let V_j be the half open vertical strip $V_j = \{(x, y) : 2^j \leq x < 2^{j+1}\}$ and let U_j be its interior. The j-th step in the modification takes place entirely in V_j. The first step of the modification of $G = G_0$ is accomplished by a slight uniform narrowing of the portion of C_1 in V_1 to obtain a domain G_1 satisfying $\overline{G_0} \cap U_1 \subset G_1$, $\{\frac{1}{2} \cdot \partial G_1\} \cap V_1 \subset G_1$, $|\varphi'(\omega_0)| < d_0 + \frac{1}{2} d_0$, and so that the new position of ω_0 is within $1/2$ of the original position. (We continue to denote by φ the self map of U corresponding to G_1.) This also changes the position of ω_k and d_k for $k > 1$, but that is of no consequence. We also use the same notation for the perturbed point and the altered angular derivative magnitude.

Proceeding by induction, we assume that $j \geq 1$ stages of the modification have been accomplished, producing a domain G_j satisfying:

(1) For each $i \geq 1$, $G_j \cap U_i$ is a disjoint union of $2i + 1$ rectangles with sides parallel to the coordinate axes,
(2) $\overline{G_{j-1}} \cap U_j \subset G_j$,
(3) $\{\frac{1}{2} \cdot \partial G_j\} \cap \bigcup_{i=1}^{j} V_i \subset G_j$,
(4) the new position of ω_i is within 2^{-j+i} of its position at the last stage, $0 \leq i \leq j-1$ (so that eventually the different positions of the ω_i form a Cauchy sequence),
(5) the new magnitude $|\varphi'(\omega_i)|$ is $< (2 - 2^{-j+i}) d_i$, $0 \leq i \leq j-1$.

For the $j+1$-st stage, we uniformly narrow the portions of $C_1, \ldots C_{j+1}$ that are in V_{j+1} by a small amount so that (1) through (5) above are satisfied when j is replaced by $j+1$. Note that the assumption that (2) holds for j allows us to have (3) satisfied for $j+1$, provided that the narrowing taking place in the $j+1$-st stage is sufficiently small. Also, (4) and (5) will hold for $j+1$ if the narrowing is small, by the continuity of the positions of the ω_i and the angular derivatives $\varphi'(\omega_i)$ under this modification.

In the limit we obtain a domain G_∞ that looks like the original one, but now with tentacles that gently narrow to the original dimensions. Moreover, $\frac{1}{2} \cdot \partial G_\infty \subset G_\infty$, and the ends of the tentacles correspond to points ω_k on the unit circle, for which $|\varphi'(\omega_k)| < 2d_k < \infty$ for each k. Then $\varphi_2'(\omega_0) = \varphi'(\varphi(\omega_0))\varphi'(\omega_0) = \varphi'(\omega_1)\varphi'(\omega_0)$, and by induction $\varphi_n'(\omega_0) = \varphi'(\omega_{n-1}) \cdots \varphi'(\omega_0)$. Hence $|\varphi_n'(\omega_0)| < 2^n d_{n-1} \cdots d_0 < \infty$, and so, for each n, C_φ^n fails to be compact by the univalent compactness theorem in §1.5.

The passage to the limit requires the fact that if $\{\varphi_j\}$ is a sequence of holomorphic self-maps of U that converges uniformly on compact subsets of U to a map φ, and if $\{\eta_j\}$ is a sequence of points on ∂U with $\eta_j \to \eta$ and $|\varphi'(\eta_j)| < M$ for each n, then φ has an angular derivative at η with $|\varphi'(\eta)| < M$. This follows from the Julia-Carathéodory Theorem (in our application the sequence $\{\eta_j\}$ represents the position of a particular ω_k at the j-th stage of modification). □

4. Proof of the Second Main Theorem

Recall that a plane domain G is *strictly starlike* if $tw \in G$ for each $w \in \overline{G}$ and $0 < t < 1$. This section treats univalent maps φ whose Königs functions take the unit disc onto strictly starlike domains. We begin by observing that unless there is considerable circular symmetry, the derivative of such a map must have "rational argument."

4.1. Lemma. *Suppose G is an unbounded strictly starlike domain and $\lambda G \subset G$ for some complex number $\lambda \neq 0$. If G is not the whole plane, then $\lambda^n > 0$ for some positive integer n.*

Proof. Choose a sequence $\{w_n\}$ in G with $|w_n| \to \infty$. By passing to a subsequence, if necessary, we may assume that the sequence $\{w_n/|w_n|\}$ converges to a complex number ω of modulus one. Let

$$L = \{r\omega : 0 \leq r < \infty\},$$

an entire ray. We claim that $L \subset \overline{G}$. To see this, fix $0 < r < \infty$. For all sufficiently large n we have $r/|w_n| < 1$, so the strict starlikeness of G insures that $rw_n/|w_n| \in G$. Thus $r\omega = \lim rw_n/|w_n| \in \overline{G}$, as desired.

Since G is invariant under multiplication by λ, so is $\overline{G}$, hence $\lambda^n L \subset \overline{G}$ for every positive integer n. Thus if the argument of λ is an irrational multiple of π, then $\overline{G}$ contains a dense set of rotates of L whereupon G is the entire complex plane. Since we are assuming this does not happen, some power of λ must be positive. □

Now we return to our standard setup: φ is a univalent self-map of U with a fixed point $p \in U$, $\lambda = \varphi'(p)$, σ is the Königs function of φ, and $G = \sigma(U)$. Recall that Schröder's equation forces $\sigma(p) = 0$, so $0 \in G$. Before getting to the heart of the proof of the Second Main Theorem, we record some standard facts about the boundary behavior of σ.

4.2. Lemma. *(a) σ has a radial limit (possibly ∞) at almost every point of ∂U. These limiting values cannot be constant on any subset of ∂U of positive measure.*

(b) If $\gamma : [0, \infty) \to G$ is a curve that tends to ∞ as $t \to \infty$, then there is a point $\omega \in \partial U$ such that

$$\lim_{t\to\infty} \sigma^{-1}(\gamma(t)) = \omega.$$

Proof. (a) This follows immediately from the fact that every univalent function on the unit disc belongs to the Hardy space H^p, for all $p \in (0, 1/2)$ ([Dur, Th. 3.16, page 50]). In our strictly starlike situation, a more elementary argument suffices. For now G cannot be dense in the plane, so if w_0 is any point at the center of a disc that does not intersect G, then $(\sigma(w) - \sigma(w_0))^{-1}$ is a bounded analytic function, and therefore has the desired properties, which it passes on to σ.

(b) Let $\Gamma = \sigma^{-1} \circ \gamma$, so Γ parameterizes a curve in U. It follows from the continuity of σ that $|\Gamma(t)| \to 1$ as $t \to 1-$. Suppose $\Gamma(t)$ does not tend to a single point on ∂U. Then the intersection of its plane closure with ∂U is a connected subset of ∂U that contains at least two points, and therefore contains one of the non-trivial open arcs between these points. Call this arc I. By part (a), σ has a radial limit at almost every point of I. But the radius to any point of I intersects Γ infinitely often, and σ tends to ∞ on Γ. Therefore σ has radial limit ∞ a.e. on I, which contradicts the conclusion of part (a). Thus Γ tends to exactly one point on ∂U, as desired. □.

We can now begin the proof of the Second Main Theorem. In view of Lemma 4.1 it is enough to assume that $0 < \lambda < 1$, in which case $n = 1$ in the statement of the Theorem. As in the proof of the First Main Theorem it suffices to prove (c) $\implies$ (a).

Suppose C_φ is not compact on H^2, so that φ has an angular derivative at some point of ∂U. We may, without loss of generality, assume this point is 1. Our goal is to show that $g = \sigma(U)$ contains a sector. Now φ has a radial limit $\varphi(1) \in \partial U$. If we can show that

$\varphi(1) = 1$, then an appeal to Proposition 3.3, along with some elementary geometry, will finish the proof. Here are the details.

First, observe that Lemma 3.5 insures $\sigma(r) \to \infty$ as $r \to 1$. Let $\theta(r)$ be a continuous determination of the argument of $\sigma(r)$. As $r \nearrow 1$ either $\lim_{r\to 1-} \theta(r)$ exists, or it does not. If it does not, the argument of the ray from the origin to $\sigma(r)$ swings between distinct upper and lower limits α and β as r tends to 1. By strict starlikeness, each of these rays lies entirely in G, so G must contain the interior of the sector bounded by the rays of arguments α and β. By the remarks following Proposition 3.3, $\sigma \notin H^p$ for some $p < \infty$.

Suppose, therefore, that the limit exists:

$$\lim_{r\to 1-} \theta(r) = \alpha.$$

The argument used to prove Lemma 4.1 shows that the entire ray

$$L = \{te^{i\alpha} : t \geq 0\}$$

lies in $\overline{G}$. Since $0 < \lambda < 1$, and G is strictly starlike, we have $L = \lambda L \subset G$. Let $\Gamma = \sigma^{-1}(L)$. Then Γ is a curve in U which by Lemma 4.2 (b) ends at a point on the unit circle:

$$\lim_{t\to\infty} \sigma^{-1}(te^{i\alpha}) = \eta \in \partial U.$$

Now $\sigma(\Gamma) = L = \lambda L = \lambda\sigma(\Gamma) = \sigma(\varphi(\Gamma))$, hence $\varphi(\Gamma) = \Gamma$. Thus $\varphi(z)$ tends to η as z tends to η along Γ, so by Lindelöf's theorem the same is true as z tends to η radially. That is, $\varphi(\eta) = \eta$.

So far we know that the angular derivative of φ exists at 1, and have just shown that φ has a boundary fixed point at $\eta = \varphi(1)$. To apply Proposition 3.3 we need only show that $\eta = 1$. With this in mind, let C denote the image of the unit interval under σ. As we are assuming that $\lim_{r\to 1-} \theta(r) = \alpha$, it may happen that C and L intersect at infinitely many points, which tend to ∞. Then back in the disc, Γ and the unit interval must intersect at the preimage points, an infinite sequence that tends to the boundary. Thus the interval and Γ share the same endpoint on ∂U, as desired.

In the remaining case, Γ has some last point r_0 of intersection with the unit interval. Let Γ_0 be the part of Γ from r_0 to η. Let L_0 be the part of L from $\sigma(r_0)$ to ∞, and let $C_0 = \sigma([r_0, 1))$. Let V be the Jordan subregion of U bounded by the interval $[r_0, 1)$, the curve Γ_0, and, in case $\eta \neq 1$ (which we are trying to show does not happen), the arc on ∂U from η to 1. Let W be the Jordan sub-region of the Riemann Sphere that is bounded by the L_0 and C_0 and contains $\sigma(V)$.

We claim that $W = \sigma(V)$. Suppose otherwise. Then W contains some point b not in G (but necessarily on ∂G). Because G is strictly starlike, it must also omit all of the line $L_b = \{\rho b : \rho \geq 1\}$. But $\lim_{r\to 1-} \theta(r) = \alpha$, so C_0 has to intersect L_b (see Figure 6 below), and this contradicts the fact that $C_0 \subset G$. The claim is established.

Thus σ takes the Jordan region V univalently onto the Jordan region W, and it follows from Carathéodory's theorem [Rud, Sections 14.19 and 14.20, pp. 310–311] that σ extends to a homeomorphism of boundaries (where ∞ is regarded as a boundary point of W). But $\sigma(\eta) = \sigma(1) = \infty$, so we must have $\eta = 1$.

So far we have established that φ has an angular derivative at the boundary fixed point $+1$, so by Proposition 3.3, G contains a twisted sector about the σ-image of the unit interval, i.e. there exists $\varepsilon > 0$ such that

$$\delta_G(\sigma(r)) \geq \varepsilon|\sigma(r)| \qquad 0 \leq r < 1.$$

Now we are assuming that $\arg\sigma(r) \to \alpha$, and that $L = \{w : \arg w = \alpha\}$. A straightforward estimate shows that because of this,

$$\delta_G(w) \geq \frac{1}{2}\varepsilon|w|$$

for all sufficiently large $w \in L$. Thus G actually contains an ordinary sector about L, as we wished to show. $\square$.

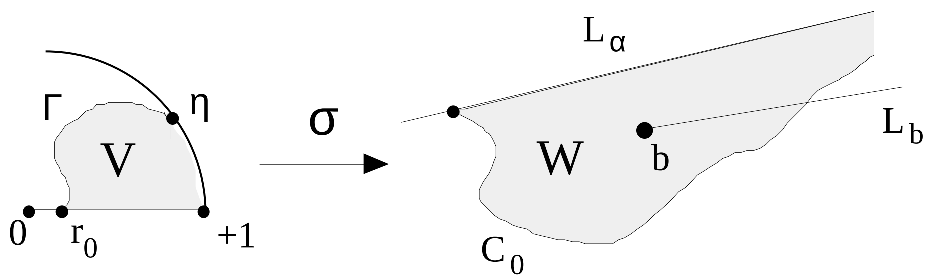

Figure 6. Regions V and W, and (nonexistent) line L_b

5. When is the Königs function in $BMOA$?

By $BMOA$ we mean the collection of functions in H^2 with boundary function of bounded mean oscillation (see [Garn, Chapter VI] for background material). Our present interest in $BMOA$ stems from two important phenomena:

(1) $H^\infty \subsetneq BMOA \subsetneq \bigcap_{p<\infty} H^p$. Moreover theorems that assert membership in H^p for every finite p frequently have $BMOA$ improvements.
(2) A univalent map of U belongs to $BMOA$ if and only if its image does not get too wide, that is, if and only if $\sup_{z\in G}\delta_G(z) < \infty$ ([Pomm2] or [Ste-St]).

As an illustration of (2) above, note that the example of section 1.6 has Königs function in $BMOA$. In this vein, recall that our initial motivation came from the fact that if a self-map of the disc induces a compact composition operator on H^2, then that map's Königs function belongs to H^p for every finite p. However our Second Main Theorem shows that the Königs function need not belong to $BMOA$. Indeed, suppose Γ is the domain that lies between the curves $y = \pm\sqrt{(x+1)}$ for $x \geq -1$, let σ denote the Riemann map of U onto G, with $\sigma(0) = 0$ and $\sigma'(0) > 0$, and define φ on U by $\varphi(z) = \sigma^{-1}(\sigma(z)/2)$. Clearly $\sigma \notin BMOA$, yet because G is strictly starlike and contains no sector, C_φ is compact.

On the other hand, the Königs function of each of the lens maps of §1.5 does lie in $BMOA$. In this section we use the lens maps to prove a comparison principle that helps predict the "$BMOA$-behavior" of σ. A surprising byproduct of our work is the fact that membership of the Königs function in $BMOA$ does not depend solely on the behavior of φ near the boundary. We will produce two univalent self-maps of the disc whose images coincide near the boundary, yet the Königs function of one will belong to $BMOA$, while that of the other will not.

The key to our analysis is the *Julia-Carathéodory ratio* of φ, defined for $z \in U$ by

$$\mathcal{J}_\varphi(z) = \frac{|\phi'(z)|(1-|z|^2)}{1-|\varphi(z)|^2}. \tag{5.1}$$

This is sometimes called the *hyperbolic change of scale* [Bear-C]; our terminology comes from the Julia-Carathéodory theorem (§2.3), which implies that if φ has an angular derivative at a point $\omega \in U$, then $\mathcal{J}_\varphi(z) \to 1$ as $z \to \omega$ non-tangentially. The geometric significance of the Julia-Carathéodory ratio is explained by the next lemma, where we continue to assume that φ is a univalent self-map of U that fixes the origin and that σ is the Königs function, so $\varphi \circ \sigma = \lambda\sigma$ (Schröder's equation), where $\lambda = \varphi'(0)$. Recall from §2.6 the hyperbolic density of $G = \sigma(U)$, given by

$$h_G(\sigma(z)) = \frac{2}{(1-|z|^2)|\sigma'(z)|}. \tag{5.2}$$

Our first result establishes the connection between the Julia-Carathéodory ratio of φ and the hyperbolic density on G.

5.1. Lemma. *For each $z \in U$,*

$$\mathcal{J}_\varphi(z) = |\lambda| \frac{h_G(\lambda\sigma(z))}{h_G(\sigma(z))}.$$

Proof. Upon differentiating both sides of Schröder's equation we obtain

$$\sigma'(\varphi(z))\varphi'(z) = \lambda\sigma'(z) \tag{5.3}$$

for each $z \in U$. Thus

$$\begin{aligned} \frac{h_G(\lambda\sigma(z))}{h_G(\sigma(z))} &= \frac{h_G(\sigma(\varphi(z)))}{h_G(\sigma(z))} && \text{(Schröder's equation)} \\ &= \frac{|\sigma'(z)|(1-|z|^2)}{|\sigma'(\varphi(z))|(1-|\varphi(z)|^2)} && \text{(by(5.2))} \\ &= \frac{|\varphi'(z)||\sigma'(z)|(1-|z|^2)}{|\lambda||\sigma'(z)|(1-|\varphi(z)|^2)} && \text{(by (5.3))} \\ &= |\lambda|^{-1}\mathcal{J}_\varphi(z). && \square \end{aligned}$$

Of particular importance is the behavior of $\mathcal{J}_\varphi$ for the "lens map" φ_α $(0 < \alpha < 1)$.

5.2. Corollary. *Suppose* $0 < \alpha < 1$. *Then* $\mathcal{J}_{\varphi_\alpha}(x) = \alpha$ *for each* $-1 < x < 1$.

Proof. We have $\lambda = \varphi_\alpha'(0) = \alpha > 0$ here, and $G = \sigma(U) = \{|\Im w| < \pi/2\}$. Since real translations are hyperbolic isometries of G, the hyperbolic density is unaffected by each such mapping. In other words, h_G is constant on horizontal lines. Since the real line is taken into itself upon multiplication by α, the result follows from Lemma 5.1. □

Of course this result could also have been verified by direct computation. It will be applied in concert with the next theorem, the main technical result of this section, which shows explicitly how the Julia-Carathéodory ratio is connected with the geometry of G.

5.3. Proposition. *Suppose* φ *is real on the real axis, and* $\varphi(1) = 1$ *(non-tangential limit). Then* $\lambda > 0$*, and if*

$$\liminf_{r\to 1-} \mathcal{J}_\varphi(r) > \lambda,$$

then

$$\lim_{r\to 1-} \delta_G(\sigma(r)) = \infty,$$

so $\sigma \notin BMOA$.

Proof. Since φ is univalent and real on the interval $(-1, 1)$, and $\varphi(r) \to 1$ as $r \to 1$, we see that φ is monotonically increasing on that interval. Thus $\lambda = \varphi'(0) > 0$. By Schröder's equation, σ, which is also univalent, is real on $(-1, 1)$, and since $\sigma'(0) \neq 0$ it is monotone increasing there. It follows readily from Schröder's equation that $\lim_{r\to 1-} \sigma(r) = \infty$.

Now our hypothesis and Lemma 5.1 combine to show that there exist positive numbers x_0 and ε such that

$$x > x_0 \implies \frac{h_G(\lambda x)}{h_G(x)} > 1 + \varepsilon.$$

Upon replacing x by $\lambda^N x$ we obtain for each positive integer N,

$$h_G(\lambda^N x) > (1 + \varepsilon) h_G(\lambda^{N-1} x) \qquad \text{for } \lambda^N x > x_0.$$

Since $\lambda < 1$ we may iterate this inequality to obtain

$$h_G(\lambda^N x) > (1 + \varepsilon)^N h_G(x) \qquad (x > \lambda^{-N} x_0). \tag{5.4}$$

Now choose N so that $(1 + \varepsilon)^N > 8$, and recall from §2.6 that the product of δ_G and h_G always lies between 1/2 and 2. Thus (5.4) and our choice of N yields

$$\delta_G(x) > 2\delta_G(\lambda^N x) \qquad (\lambda^{-N} x > x_0).$$

Upon replacing x by $\lambda^{-N} x$ in this inequality, we obtain

$$\delta_G(\lambda^{-N} x) > 2\delta_G(x) \qquad (x > x_0).$$

To finish the proof, let $\delta_0 = \min\{\delta_G(x) : x_0 \leq x < \lambda^{-N} x_0\}$. Then, for $\lambda^{-kN} x_0 \leq x < \lambda^{-(k+1)N} x_0$, we have

$$\delta_G(x) > 2\delta_G(\lambda^N x) > 4\delta_G(\lambda^{2N} x) \cdots > 2^k \delta_G(\lambda^{kN} x) \geq 2^k \delta_0.$$

Thus $\delta_G(x) \to \infty$ as $x \to \infty$, as desired. □

Proposition 5.3 provides the means for constructing the class of examples promised at the beginning of this section. In the statement below, $0 < \alpha < 1$, and φ_α is, as usual, the lens map defined earlier.

5.4. Theorem. *Suppose $\varphi = \varphi_\alpha \circ \psi$ where ψ is a univalent self-map of U with these additional properties:*

(1) *$\psi(0) = 0$ and $\psi(1) = 1$ (non-tangential limit).*
(2) *ψ is real on the real axis.*
(3) *The angular derivative of ψ exists at the boundary point 1.*
(4) *ψ is not the identity map.*

Then the composition operator C_φ is compact on H^2, but the Königs function of φ is not in $BMOA$.

Proof. We know that C_{φ_α} is compact, and the definition of φ, interpreted on the operator level, says that $C_\varphi = C_\psi C_{\varphi_\alpha}$. Thus C_φ is compact, since the product (in either order) of a bounded operator and a compact operator is compact (see [Con, Prop. 4.2, page 41], for example).

To see that $\sigma \notin BMOA$ we compute the Julia-Carathéodory ratio of φ on the interval $(-1, 1)$. Note that, as before, the univalence and symmetry assumed for φ insures that ψ is monotonically increasing on $(-1, 1)$. Applying (in this order) the Chain Rule, Corollary 5.2, and the Julia-Carathéodory theorem, we obtain for $-1 < r < 1$,

$$\mathcal{J}_\varphi(r) = \mathcal{J}_{\varphi_\alpha}(\psi(r))\mathcal{J}_\psi(r) \to \alpha \qquad \text{as } r \to 1- .$$

But $\psi'(0) < 1$ since ψ is not the identity map, hence

$$\varphi'(0) = \varphi'_\alpha(0)\psi'(0) < \varphi'_\alpha(0) = \alpha,$$

and so $\lim_{r\to 1-} \mathcal{J}_\varphi(r) > \varphi'(0)$. Thus φ satisfies the hypotheses of Proposition 5.3, so its Königs function is not in $BMOA$. □

5.5. A class of examples. Let $V_\alpha = \varphi_\alpha(U)$, where φ_α is a lens map, and take V to be any proper subregion of V_α that is symmetric about the real axis and coincides with V_α in some disc centered at the boundary point 1. For example, V could be the standard lens with two small subintervals of equal length removed from the imaginary axis, with outer endpoints on the boundary of the lens.

Let φ be the univalent map of U onto V that fixes both the origin and the point 1, and has positive derivative at the origin. Then $\psi = \varphi_\alpha^{-1} \circ \varphi$ is a univalent self-map of U that clearly obeys hypotheses (1), (2), and (4) of Theorem 5.4, and is not the identity map. Because the images of φ and φ_α coincide in a neighborhood of the point 1, ψ maps an arc of ∂U centered at 1 to another such arc. Thus ψ has an analytic extension across that arc, and in particular it has an angular derivative (in fact an ordinary complex derivative) at 1. The hypotheses of Theorem 5.4 are therefore satisfied, so the Königs function of φ is not in $BMOA$. □

5.6. Vanishing Mean Oscillation. The notion of "bounded mean oscillation" is a "big-oh" condition asserting that certain integral means involving the amount by which a function differs from its average over an interval, when taken over all subintervals of ∂U, form a bounded set. The corresponding "little-oh" condition, as the lengths of those subintervals tend to zero, is called "vanishing mean oscillation." The corresponding subspace of $BMOA$ is called $VMOA$. For our purposes it is enough to know that a univalent mapping of the unit disc onto a region G is in $VMOA$ if and only if $\delta_G(w) \to 0$ as $w \to \infty$ through G (see, as before, [Pomm2], or [Ste-St]).

It is therefore not surprising that appropriate modifications of the arguments given above lead to these companion theorems for $VMOA$:

(1) Under the hypotheses of Proposition 5.3,

$$\limsup_{r\to 1-} \mathcal{J}_\varphi(r) < |\lambda| \implies \lim_{r\to 1-} \delta_G(\sigma(r)) = 0.$$

(2) In particular, suppose φ fixes both the origin and the point 1, and $\varphi(U)$ is symmetric about the real axis, properly contains a standard lens, and coincides with that lens outside some disc $|z| < r < 1$. Then the Königs function of φ is in $VMOA$.

As an example of the situation in (2) above, take $G = \varphi(U)$ to be a standard lens with two small discs of equal radius adjoined so that their centers lie at the points where the imaginary axis intersects the boundary of the lens.

5.7. Symmetry assumptions. The pervasive assumption of symmetry about the real axis first appeared in Proposition 5.3. It can be weakened a little; the arguments will still work if we only assume that the curve $\sigma([0,1])$ is taken into itself upon multiplication by λ. However without some such extra hypotheses, the Proposition is false, as the following example shows.

Example: $\liminf_{r\to 1-} \mathcal{J}_\varphi(r) > |\varphi'(0)|$, *but* $\sigma \in BMOA$. Suppose that $\lambda \notin \mathbb{R}$ has small modulus and small argument. Let γ be the curve $\{\lambda^t \mid -\infty < t < \infty\}$, so that $\lambda\gamma = \gamma$. Now take Ω to be a simply connected domain containing 0 with the properties that $\gamma \cap \{z : |z| \geq R\} \subset \partial\Omega$, for some $R > 0$, and that $\delta_\Omega(w)$ is bounded. Thus Ω can be visualized as the union of a disk centered at the origin and an unbounded spiraling strip with γ an edge of the strip. The conformal mapping $\sigma : U \to \Omega$ with $\sigma(0) = 0$ and $\sigma([0,1))$ unbounded is in $BMOA$, since $\delta_\Omega(w)$ is bounded. Furthermore, since $\gamma \cap \{z : |z| \geq R\} \subset \partial\Omega$ and γ is invariant under multiplication by λ, it follows that

$$\frac{\delta_\Omega(\sigma(r))}{\delta_\Omega(\lambda\sigma(r))} \geq C\frac{1}{|\lambda|}, \qquad r \geq r_0,$$

for some $r_0 \in (0,1)$ and $C > 0$. Using (2) as in the proof of Lemma 5.1, we see that

$$\frac{h_\Omega(\lambda\sigma(r))}{h_\Omega(\sigma(r))} \geq \frac{1}{4} \cdot \frac{\delta_\Omega(\sigma(r))}{\delta_\Omega(\lambda\sigma(r))}.$$

Thus if $|\lambda|$ is sufficiently small, we will achieve the promised inequality on the Julia-Carathéodory ratio.

It may be objected that this example is special in that a large part of the boundary of Ω is taken into itself upon multiplication by λ, but it is clear that a small perturbation of Ω will produce a domain G for which $\lambda \cdot \partial\Omega \subset \Omega$, and for which all the other features of the original example are preserved.

5.8. Higher orders of compactness. The examples produced in §5.5 are actually "very compact" in that they belong to all the *Schatten p-classes* for $p < \infty$. (see [Sh-Tay, page 496]). Thus, even in the presence of higher orders of compactness, the Königs function need not belong to $BMOA$. On the other hand, the condition $\|\varphi\|_\infty < 1$, which, as we saw in §1, forces σ to be bounded, can be regarded as inducing even stronger compactness on C_φ (for example, C_φ now takes the unit ball into a subset that is compact in H^∞). It might be interesting to to see if there are reasonable intermediate "hyper-compactness" conditions, say on the approximation numbers of C_φ, that force the Königs function of φ into $BMOA$.

6. Final remarks.

We close with some comments about how our main results fare in different settings.

6.1. The Bergman spaces. The *Bergman space* A^p of the unit disc is the space of functions f holomorphic on U for which

$$\|f\|_p^p = \int |f|^p \, dA < \infty,$$

where $dA = \frac{1}{\pi} dxdy$ is normalized Lebesgue measure on U. A^2 is a Hilbert space on which composition operators act boundedly (in fact, composition operators act boundedly on H^p and A^p regardless of p), and on which the angular derivative criterion of §2.5 characterizes compactness even in the absence of univalence (see [M-S] or [Sh, §6]).

Now it is easy to check that for each $p < \infty$ the Bergman space A^p contains H^p, so the same is true of their intersections for $p < \infty$. In fact this containment of intersections is *proper*. For example, the function $\sum z^{2^n}$ belongs to the Bloch space, which is contained in every space A^p for $p < \infty$, but it belongs to *no* H^p since it has radial limits almost nowhere (see for example [Pomm3, Example 1, page 696]). Nevertheless:

Proposition. *Under the hypotheses of either of our two Main Theorems, if the Königs function of φ lies in A^p for every $p < \infty$, then some power of C_φ is compact on H^2.*

To see why this is true, we need only note that the only property of H^p needed for the proofs of either Main Theorem is the fact that for each $f \in H^p$, the modulus $|f(z)|$ is bounded by a constant multiple of a negative power of $1 - |z|$. In fact the same is true for A^p.

Lemma. *If $f \in A^p$ then for each $z \in U$,*

$$|f(z)| \le \frac{\|f\|_p}{(1 - |z|)^{2/p}}.$$

Proof. For $z \in U$, let $\Delta_z = \{w : |w - z| < 1 - |z|\}$. Then the subharmonicity of $|f|^p$ yields

$$|f(z)|^p \le \frac{1}{A(\Delta_z)} \int_{\Delta_z} |f|^p \, dA \le \frac{\|f\|_p^p}{(1 - |z|)^2}$$

as promised. $\square$

In view of the folk theorem of §1.2, there results this curious fact:

If, under the hypotheses of either of the Main Theorems, $\sigma \in \cap_{p<\infty} A^p$, *then* $\sigma \in \cap_{p<\infty} H^p$.

6.2. The Dirichlet space. The *Dirichlet space* $\mathcal{D}$ of the unit disc is the space of functions f holomorphic on U for which $f' \in A^2$. $\mathcal{D}$ is a Hilbert space in the norm $\| \ \|$ defined by

$$\|f\|^2 = |f(0)|^2 + \int_U |f'|^2 \, dA < \infty.$$

Since the integral on the right is the "multiplicity area" of the image of f, a univalent function is in $\mathcal{D}$ if and only if it takes the unit disc onto a region with finite area. If φ is a univalent self-map of the unit disc, then a simple change of variable shows that the composition operator induced by φ acts boundedly on $\mathcal{D}$. The compactness of such operators is characterized by the following lemma ([M-S], Proposition 5.1, p. 892, with $\alpha = 0$).

Lemma. *If* φ *is a univalent mapping of the disc into the disc, then* C_φ *is compact on* $\mathcal{D}$ *if and only if*

$$\lim_{\epsilon \to 0+} \sup_{\zeta \in \partial U} A(\varphi(\{z \in U \mid |z-\zeta| < \epsilon\}))/\epsilon^2 = 0.$$

$\mathcal{D}$ is a proper subset of H^2, and one might hope that analogs of our two Main Theorems would hold for $\mathcal{D}$ as well. In particular, does $\sigma^n \in \mathcal{D}$ for all integers $n \geq 1$ (the analog of $\sigma \in H^p$ for all $p < \infty$) imply that C_φ is compact on $\mathcal{D}$? The following example shows that the answer is no.

Example. Let

$$G = \bigcup_{k=0}^{\infty} R_k,$$

where R_0 is the unit disc, and for each $k \geq 1$,

$$R_k = (0, 4^k) \times (-\exp(-4^k), \exp(-4^k)).$$

Define σ to be the conformal map from U to G with $\sigma(0) = 0$ and $\sigma'(0) > 0$, and define φ by $\varphi(z) = \sigma^{-1}(\sigma(z)/2)$, $z \in U$. It is not difficult to see that, for each integer $n \geq 1$, $\sigma^n \in \mathcal{D}$, since the restriction of σ^n to the intersection of U with a sufficiently small disc centered at 1 is univalent.

On the other hand, φ is modeled by multiplication by $1/2$ on G, and a simple estimate using the quasi-hyperbolic metric (see §2.6) shows that there is a constant $C > 0$ such that the hyperbolic distance on G satisfies

$$\frac{1}{C} \leq \rho_G((x,0), \partial(\tfrac{1}{2}G)) \leq C,$$

for each integer $k \geq 1$ and x such that $4^k \leq x \leq 2 \cdot 4^k$. By conformal invariance of the hyperbolic metric, on U this becomes

$$\frac{1}{C} \leq \rho_U(r, \partial(\varphi(U))) \leq C, \qquad 4^k \leq \varphi(r) \leq 2 \cdot 4^k.$$

For each integer $k \geq 1$, define $r_k \in (0,1)$ by $\varphi(r_k) = 4^k$, so that $r_k \to 1$. We have seen that $\varphi(U)$ contains a hyperbolic disc centered at r_k of radius $1/C$, so it contains an Euclidean disc of radius $(1-r_k)/C_1$ and center r_k, where C_1 is comparable to C. Since C_1 is independent of k and $r_k \to 1$, the Lemma characterizing compact composition operators on $\mathcal{D}$ shows that C_φ is not compact.

References

[Ahl] L.V. Ahlfors, *Conformal Invariants*, McGraw-Hill, New York, 1973.

[BaPomm] I.N. Baker and Ch. Pommerenke, *On the iteration of analytic functions in a half-plane II*, J. London Math. Soc. (2) **20** (1979), 255 – 258.

[Bear-C] Beardon, A. F. and Carne, T. K., *A strengthening of the Schwarz-Pick inequality*, Amer. Math. Monthly **99** (1992), 216 – 217.

[BoS] Bourdon, P.S. and Shapiro, J.H., *Cyclic phenomena for composition operators*, Preprint.

[Car] Caratheodory, C., *Theory of Functions Volume Two*, Verlag Birkhäuser AG., Basel, 1950.

[CarCow] Carroll, T.F. and Cowen, C.C., *Compact composition operators not in the Schatten classes*, J. Operator Theory **26** (1991), 109 – 120.

[C-S] J. Caughren and H.J. Schwartz, *Spectra of compact composition operators*, Proc. Amer. Math. Soc. **51** (1975), 127 – 130.

[Cow] Cowen, C.C, *Composition operators on Hilbert spaces of analytic functions: a status report*, Proc. Sympos. Pure Math. **51** (1990), 131–145, In: Operator Theory: Operator Algebras and applications, Part I (W.B. Arveson and R.G. Douglas, eds) Amer. Math.Soc., Providence, RI..

[Cow2] Cowen, C.C, *Iteration and the solution of functional equations for functions analytic in the unit disc*, Trans. Amer. Math. Soc. **265** (1981), 69 – 95.

[Cow3] Cowen, C.C, *Composition operators on H^2*, J. Operator Theory. **9** (1983), 77 – 106.

[CowKr] Cowen, C.C. and Kreite, T.L. III, *Subnormality and composition operators on H^2*, J. Funct. Anal. **81** (1988), 298 – 319.

[Con] Conway, J.B., *A Course in Functional Analysis*, Springer-Verlag,, New York, 1985.

[Dur] Duren, P.L, *Theory of H^p Spaces*, Academic Press, New York, 1970.

[Garn] Garnett, J. B., *Bounded Analytic Functions*, Academic Press, New York, 1981.

[Ge-Pal] Gehring, F. W. and Palka, B. P., *Quasiconformally homogeneous domains*, J. Analyse Math. **30** (1976), 172–199.

[Ko] Königs , G., *Recherches sur les intégrales de certaines équations fonctionnelles*, Ann. Sci. École Norm. Sup., (troisieme série), supplément **1** (1884), 3–41.

[Lit] Littlewood, J.E., *On inequalities in the theory of functions*, Proc. London Math. Soc. **23** (1925), 481–519.

[M-S] MacCluer, B.D. and Shapiro, J.H., *Angular derivatives and compact composition operators on the Hardy and Bergman spaces*, Canad. J. Math. **38** (1986), 878–906.

[Na-Pal] Näkki, R. and Palka, B. P., *Asymptotic values and Hölder continuity of quasiconformal mappings*, J. Anayse Math. **48** (1987), 167–178.

[Pomm1] Pommerenke, Ch., *Univalent Functions*, Vandenhoeck and Ruprecht, Göttingen, 1975.

[Pomm2] Pommerenke, Ch., *Schlichte Funktionen und analytische Funtionen von beschränkter mittler Ozillation*, Comm. Math. Helv **52** (1977), 591 – 602.

[Pomm3] Pommerenke, Ch., *On Bloch functions*, J. London Math. Soc (2) **2** (1970), 689–695.

[Pomm4] Pommerenke, Ch., *On the iteration of analytic functions in a half-plane I*, J. London Math Soc. (2) **19** (1979), 439 – 447.

[Rud] Rudin, W., *Real and Complex Analysis*, second ed., McGraw-Hill, New York, 1974.

[Sch] E. Schröder, *Über iterierte Funktionen*, Math. Ann. **2** (1870), 317–365.

[Sh-Tay] Shapiro, J.H and Taylor, P.D, *Compact, Nuclear, and Hilbert-Schmidt composition operators on* H^p, Indiana Univ. Math. J. **23** (1973), 471–496.

[Sh] Shapiro, J.H., *The essential norm of a composition operator*, Annals of Math. **127** (1987), 375–404.

[Sm-Ste] Smith, W., and Stegenga, D. A., *Hölder domainsand Poincaré domains*, Trans. Amer. Math. Soc. **319** (1990), 67–100.

[Ste-St] Stegenga, D. A. and Stephenson, K, *A geometric characterization of analytic functions with bounded mean oscillation*, J. London Math. Soc. (2) **24** (1981), 243–254.

[Val] Valiron, G, *Sur l'iteration des fonctions holomorphes dans un demi-plan*, Bull des Sci. Math. (2) **55** (1931), 105 – 128.

Dated: 12/16/1993

Department of Mathematics, Michigan State University, East Lansing, MI 48824
E-mail address: shapiro@math.msu.edu

Department of Mathematics, University of Hawaii, Honolulu, Hawaii 96822
E-mail address: wayne@math.hawaii.edu and dave@math.hawaii.edu