LECTURES ON COMPOSITION OPERATORS AND ANALYTIC FUNCTION THEORY

JOEL H. SHAPIRO

1. Invertibility and the Schwarz Lemma

1.1. **Introduction.** At first we will work in $H(U)$, the collection of all complex-valued functions that are holomorphic (i.e., analytic) on on the unit disc $U = \{z \in \mathbf{C} : |z| < 1\}$ of the complex plane. Because pointwise sums and products of analytic functions are again analytic, these pointwise operations make $H(U)$ into an *algebra,* and in particular a *vector space,* over the field of complex numbers.

Suppose φ is a holomorphic *self-map of* U*,* i.e. φ is holomorphic on U and $\varphi(U) \subset U$. From now on we always use the symbol φ to represent—usually without further comment—a holomorphic self-map of U. For $f \in H(U)$ and $z \in U$ define $C_\varphi f : U \to \mathbf{C}$ by:

$$(C_\varphi f)(z) = f(\varphi(z)).$$

Since compositions of analytic functions are analytic when the domains and ranges match up correctly (which they do here), we see that $C_\varphi f \in H(U)$, so we have defined a mapping $C_\varphi : H(U) \to H(U)$ that's called the *composition operator* induced by φ.

It is easy to check that C_φ is a linear transformation on $H(U)$. The point of studying composition operators is to understand how the properties of the analytic function φ influence those of the linear map C_φ, and *vice versa.*

We start with what is perhaps the most basic question you can ask about any linear transformation: *when is it invertible?* In other words, when is it one-to-one and onto? For composition operators, one-to-one is never an issue:

1.2. **Theorem.** *If φ is not constant than C_φ is one-to-one on $H(U)$.*

Proof. Suppose that for some f and g in $H(U)$ we have $C_\varphi f = C_\varphi g$, i.e. that $f(\varphi(z)) = g(\varphi(z))$ for all $z \in U$. Then $f \equiv g$ on $\varphi(U)$. But nonconstant holomorphic functions are *open mappings*, so $f \equiv g$ on a nonempty open subset of U. Thus f and g agree on a subset of U having a limit point in U, so by the *uniqueness theorem* $f \equiv g$ on U. Thus C_φ is one-to-one. □

That was easy, but note that it required two important results about analytic functions—the open mapping theorem and the uniqueness theorem.

Date: May 28, 1998.

1.3. **Exercise.** Let $C(U)$ denote the collection of complex valued continuous functions on U. For $\psi : U \to U$ a continuous self-map of U, define C_ψ on $C(U)$ in the obvious way. Determine those ψ's for which C_ψ is one-to-one.

In order to characterize the composition operators that are invertible on $H(U)$, we have to work a little harder.

1.4. **Theorem.** *C_φ is invertible on $H(U)$ if and only if φ maps U one-to-one onto itself.*

Proof. It's clear that if φ maps U one-to-one onto itself then C_φ is invertible, since then the inverse is $C_{\varphi^{-1}}$.

The converse takes more work. We will show that if φ is not an automorphism, then the range C_φ is not $H(U)$. If φ is not one-to-one this is easy: there exist distinct points a and b in U with $\varphi(a) = \varphi(b)$, hence for every $f \in H(U)$ the image function $C_\varphi f = f \circ \varphi$ takes the same value at a that it takes at b. In particular, the identity map $f(z) \equiv z$ cannot belong to $\operatorname{ran} C_\varphi$, so C_φ is not invertible.

One case remains: suppose $\varphi(U) \neq U$. Then we can choose $w \in U \backslash \varphi(U)$ and consider the functions

$$f(z) = \frac{1}{z - w}, \quad \text{and} \quad g(z) = \frac{1}{\varphi(z) - w} \qquad (z \in U).$$

Because $w \notin \varphi(U)$, we see that f is holomorphic on $\varphi(U)$, and g is holomorphic on U.

I claim that $g \notin \operatorname{ran} C_\varphi$. Suppose otherwise. Then there exists $h \in H(U)$ such that $h(\varphi(z)) = g(z) = f(\varphi(z))$ for every $z \in U \backslash \{w\}$. Thus $h = f$ on a nonempty open subset of $U \backslash \{w\}$, and so $h = f$ on all of $U \backslash \{w\}$, again by the identity theorem. Thus f must be bounded in a neighborhood of w, which it obviously is not. This shows that $\operatorname{ran} C_\varphi \neq H(U)$, as desired. □

1.5. **Exercise.** Show that if ψ is a continuous self-map of the unit disc, then C_ψ is an isomorphism of $C(U)$ if and only if ψ is a homeomorphism of U onto itself.

Conformal automorphisms of U. One-to-one holomorphic mappings are usually called *univalent.* If a holomorphic self-map of the unit disc is univalent and maps U *onto* itself, we call it a *conformal automorphism* (or just an *automorphism* of U. Theorem 1.4 can now be rephrased:

> *C_φ is invertible on $H(U)$ if and only if φ is a conformal automorphism of U.*

To make this result meaningful, however, we have to *characterize* conformal automorphisms in some concrete way. A good beginning involves finding a large class of examples.

For $p \in U$ define

$$\alpha_p(z) \stackrel{\text{def}}{=} \frac{p - z}{1 - \bar{p}z} \qquad (z \in U). \tag{1.1}$$

Because the denominator is zero only at $z = 1/\bar{p}$, the function α_p is holomorphic on the disc $\{|z| < 1/|p|\}$, which contains the closed unit disc. Note that α_p interchanges the origin and the point p.

1.6. **Theorem.** *α_p is a conformal automorphism of U.*

Proof. The proof hinges on the formula below, which I leave to you as an exercise.

$$1 - |\alpha_p(z)|^2 = \frac{(1-|p|^2)(1-|z|^2)}{|1-\bar{p}z|^2} \qquad (z \in \mathbf{C}\backslash\{1/\bar{p}\}). \tag{1.2}$$

The right-hand side of this equation is > 0 for every $z \in U$, and $= 0$ for every $z \in \partial U$. Thus α_p maps U into itself, and ∂U into itself.

The fact that α_p is a conformal automorphism follows from this simple exercise: *α_p is its own compositional inverse,* in other words, $\alpha_p(\alpha_p(z)) = z$. This is most easily seen by reducing the equation $w = \alpha_p(z)$ to its equivalent form $w - w\bar{p}z + z = p$, in which w and z play symmetric roles. □

1.7. **Exercise.** *If p and q are points of U then there is a conformal automorphism of U that maps p onto q.*

Clearly the conformal automorphisms of U form a group under composition. The exercise above says that this group acts *transitively* on U.

We still haven't classified *all* the conformal automorphisms of U. The largest class that comes to mind, given what we've already done, is the one you get by composing maps α_p with rotations, i.e., maps of the form $\alpha_p(\omega z)$ or $\omega\alpha_p(z)$ for $p \in U$, and then composing these maps with each other . . . . It is a remarkable fact that nothing new arises from this process.

1.8. **Theorem.** *If φ is a conformal automorphism of U then there exists $p \in U$ and $\omega, \eta \in \partial U$ such that $\varphi(z) = \omega\alpha_p(z) = \alpha_p(\eta z)$ for all $z \in U$.*

Proof. I'll prove the first equality and leave the second one as an exercise. Since φ maps U onto itself there is a point $p \in U$ such that $\varphi(p) = 0$. Thus $\varphi(\alpha_p(0)) = 0$, so if we define $\psi = \varphi \circ \alpha_p$ then ψ is a conformal automorphism of U that fixes the origin. Now the only such maps that come quickly to mind are the rotations. In fact we'll see shortly that there are no others. Granting this, there must be a complex number ω of modulus one such that

$$\omega z = \psi(z) = \varphi(\alpha_p(z)) \qquad (z \in U).$$

Upon replacing z by $\alpha_p(z)$ in this equation and using the self-inverse property of α_p, we see that $\omega\alpha_p(z) = \varphi(z)$ for each $z \in U$, as desired. □

It remains to prove that the only conformal automorphisms of U that fix the origin are the rotations. This is not a triviality—it requires one of the most important theorems in complex analysis:

1.9. **The Schwarz Lemma.** *If φ is a holomorphic self-map of U with $\varphi(0) = 0$, then*

(a) $|\varphi(z)| \leq |z|$ *for every z in U, with equality for some $z \in U$ if and only if φ is a rotation (about the origin).*
(b) $|\varphi'(0)| \leq 1$, *with equality if and only if φ is a rotation.*

Proof. Everything follows from an analysis of the function g defined on U by: $g(z) = \varphi(z)/z$ if $z \in U\backslash\{0\}$, and $\varphi'(0)$. Then g is holomorphic on U. Fix $z_0 \in U$ and consider a number r with $|z_0| < r < 1$. Note that since $|g(z)| \leq 1/r$ on the circle $\{|z| = r\}$, the maximum modulus principle shows that $|g(z_0)| \leq 1/r$. Upon letting $r \to 1-$ we see that in fact $|g(z_0)| \leq 1$, hence $|\varphi(z_0)| \leq |z_0|$.

As for equality in part (a), suppose that for some $z_0 \in U$ we had $|\varphi(z_0)| = |z_0|$. Then $|g(z_0)| = 1$ so if g were not constant, then by the maximum principle we would have $|g| > 1$ somewhere on U. Since this does not happen, g is constant, necessarily a constant of modulus one, and therefore φ is a rotation. This proves part (a) of the Schwarz Lemma.

The "inequality" part of part (b) follows immediately from the corresponding part of part (a). As for the rest, if $|\varphi'(0)| = 1$ then g attains its maximum at the origin, so as before, g has to be constant, and therefore φ is a rotation. □

1.10. **Corollary.** *If φ is a conformal automorphism of U that fixes the origin, then there exists $\omega \in \partial U$ such that $\varphi(z) = \omega z$ for every $z \in U$.*

Proof. Since φ is a holomorphic self-map of U that fixes the origin, part (b) of the Schwarz Lemma guarantees that $|\varphi'(0)| \leq 1$. But since φ is an automorphism, it has a compositional inverse ψ that also obeys the hypotheses of the Schwarz Lemma, hence $|\psi'(0)| < 1$. By the chain rule, $\varphi'(0)\psi'(0) = 1$, hence $|\varphi'(0)| = 1$, and so by the "equality part" of part (b) of the Schwarz Lemma, φ is a rotation. □

With this, the proof of Theorem 1.8 is complete.

The Schwarz Lemma will prove to be the single most valuable result in this course. Here is another application, seemingly not related to invertibility, which we will need in the next section. For this one we use the notation

$$\varphi_n = \varphi \circ \varphi \circ \ldots \circ \varphi \qquad (n \text{ times}).$$

to denote the *n-th iterate* of φ for n a positive integer.

1.11. **Theorem.** *If φ is a holomorphic self-map of U with $\varphi(0) = 0$, and φ is not a rotation, then $\varphi_n \to 0$ uniformly on compact subsets of U.*

Proof. Fix is a compact subset K of U, and fix $0 < r < 1$ so that $K \subset \{|z| < r\}$. By the Schwarz Lemma we know that $|\varphi(z)| < r$ on the circle $\gamma_r = \{|z| = r\}$, the strict inequality coming from the fact that φ is not a rotation. Since $|\varphi|$ is continuous and

γ_r compact, we know that $|\varphi|$ attains its maximum $M(r)$ on γ_r, so also $M(r) < r$. This means that the function ψ, defined on U by

$$\psi(w) = \frac{1}{M(r)}\varphi(rw) \qquad (w \in U),$$

is a holomorphic self-map of U that fixes the origin. So by the Schwarz Lemma, $\psi(w) \leq |w|$ for all $w \in U$ (in fact, there is strict inequality since ψ is not a rotation either). Translating this estimate into the language of φ we obtain

$$(1.3) \qquad |\varphi(z)| \leq \delta|z| \qquad (|z| < r)$$

where $\delta = M(r)/r < 1$. Since φ takes the disc $\{|z| < r\}$ into itself, the above estimate can be iterated, resulting in

$$(1.4) \qquad |\varphi_n(z)| \leq \delta^n|z| \qquad (|z| < r, n = 1, 2, \dots).$$

In particular, $\max_{z\in K} |\varphi_n(z)| \leq \delta^n \max_{z\in K} |z|$, hence $\varphi_n \to 0$ uniformly on K. □

1.12. **Exercise.** Show that if φ is a holomorphic self-map of U that fixes a point $p \in U$, and is not a conformal automorphism, then $\varphi_n \to p$ uniformly on compact subsets of U (the above theorem is the case $p = 0$).

Let's close this section with a few more exercises that illustrate further uses of the Schwarz lemma.

1.13. **Exercise.** Show that if φ is a holomorphic self-map of U then for every $p \in U$,

$$|\varphi'(p)| \leq \frac{1 - |\varphi(p)|}{1 - |p|}$$

with equality if and only if there is a complex number of modulus one such that $\varphi(z) = \omega\alpha_p(z)$ for all $z \in U$. (Hint: Let $q = \varphi(p)$ and consider $\alpha_q \circ \varphi \circ \alpha_p$).

1.14. **Exercise.** Show that if a holomorphic self-map φ of U fixes two distinct points of U, then φ is the identity map. (Hint: Consider first the case where one of the fixed points is the origin.)

1.15. **Exercise.** Show that if φ is a holomorphic self-map of U that fixes a point $p \in U$, then $|\varphi'(p)| \leq 1$, with equality if and only if φ is an automorphism.

2. Eigenvalues and Schröder's equation

Introduction. For linear transformations the next most fundamental topic after invertibility is *eigenvalues.* The eigenvalue equation for a composition operator C_φ on $H(U)$ is $C_\varphi f = \lambda f$, or more classically:

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

a functional equation which, given φ, is to be solved for f and λ. Equation (2.1) is called *Schröder's equation;* it was first studied by Ernst Schröder during the 1860's in connection with iteration of holomorphic functions.

To understand this connection, suppose we are lucky enough to have solved Schröder's equation with a *univalent* eigenfunction f. Then $\varphi = f^{-1} \circ M_\lambda \circ f$, where M_λ is the map of multiplication by the complex number λ, acting on $f(U)$ (Schröder's equation insures that M_λ takes $f(U)$ into itself). In other words, the action of φ on U is "modeled" by the action of the much simpler mapping M_λ on the more complicated domain $f(U)$. For obvious reasons, then, Schröder's equation is said to "linearize" the action of φ.

The connection with iteration comes from the fact that

$$\varphi_n = f^{-1} \circ M_{\lambda^n} \circ f \qquad (n = 1, 2, \dots). \tag{2.2}$$

For example, if you wish to study the φ-*orbit* $\{\varphi_n(z_0)\}$ of a point $z_0 \in U$, all you have to do is pull back the easily visualized M_λ-orbit $\{\lambda^n f(z_0)\}$ via the map f.

It is rare that Schröder's equation can be solved explicitly. Here are two examples where you can do it. The first one has a fixed point at the origin, and the second has no fixed point in U. Note that the properties of the solutions are very different.

2.1. **Exercise.**

(a) Suppose ω is a complex number of modulus 1, and set $\varphi(z) = \omega z$. Show that the numbers $1, \omega, \omega^2, \dots$ are all eigenvalues of C_φ. Show eigenvalues have multiplicity one if ω is not a root of unity, and infinite multiplicity otherwise.

(b) Let $\varphi(z) = \frac{1+z}{2}$. Show that every nonzero complex number is an eigenvalue of C_φ. Do these eigenvalues have finite multiplicity? (Hint: Consider the functions $f_\alpha(z) = (1-z)^\alpha$ for $\alpha \in \mathbf{C}$.)

Schröder never succeeded in getting a solution to his equation for any significant classes of maps φ, and was forced to settle for what he called the "opposite approach" wherein one constructs examples by choosing a mapping f that takes the unit disc univalently onto a domain that is invariant under multiplication by a complex number λ of modulus < 1, and defining φ on U by equation (2.2). Here are some examples for you to work out.

2.2. **Exercise.** Let G denote the half-plane $\{w \in \mathbf{C} : \operatorname{Re} w > -1\}$, and let $f(z) = \frac{2z}{1-z} = \frac{1+z}{1-z} - 1$. Prove that f maps U onto G, taking the unit circle to the line $\operatorname{Re} z = -1$ (with 1 going to ∞). Show that the multiplication map M_λ takes Π into itself if and only if $\lambda > 0$. Thus for each positive λ the right-hand side of equation

(2.2) defines a holomorphic self-map φ_λ of U. Show that $\varphi_\lambda(0) = 0$, and that φ is not an automorphism. Determine the orbit of an arbitrary point $z \in U$ under the action of φ. Show that the numbers λ^n are eigenvalues of C_{φ_λ} $(n = 0, 1, 2, \dots)$. Find the corresponding eigenfunctions.

2.3. **Exercise.** Same setup as in the previous exercise, but replace G by the right half-plane $\Pi = \{w \in \mathbf{C} : \operatorname{Re} w > 0\}$, and set $f(z) = \frac{1+z}{1-z}$. This time show that f maps U onto Π, and that for each $\lambda > 0$: φ_λ is an automorphism of U that fixes the points ± 1 and no others. In particular, φ_λ has no fixed point in U. Show that the numbers λ^n are eigenvalues of C_{φ_λ} $(n = 0, 1, 2, \dots)$. Find the corresponding eigenfunctions.

Real progress on the theory of Schröder's equation had to wait until 1884. In that year Gabriel Koenigs published his work on solutions of Schröder's equation for the class of holomorphic self-maps φ that are *not* conformal automorphisms, that fix a point of $p \in U$ and for which $\varphi'(p) \neq 0$ (so that φ is, in particular, nonconstant). Let us call these mappings *Koenigs maps.* Note that by Exercise 1.15 we have $0 < |\varphi'(p)| < 1$ for each Koenigs map φ with fixed point at $p \in U$.

Here is the formal statement of the result that Koenigs proved.

2.4. **Theorem of Koenigs.** *If φ is a Koenigs map, then the eigenvalues of C_φ are precisely the numbers $\varphi'(p)^n$ for $n = 0, 1, 2, \dots$. Moreover these are the only eigenvalues, and they all have multiplicity one. If φ is univalent, then so is the eigenfunction σ corresponding to the eigenvalue $\varphi'(p)$.*

The "multiplicity one" property of eigenvalues justifies our reference to "the" eigenfunction corresponding to the eigenvalue $\varphi'(p)$. Note that it is clear that once you have an eigenvector σ for $\varphi'(p)$, then for any positive integer n, $\varphi'(p)^n$ is an eigenvalue of C_φ with eigenvector σ^n. Multiplicity 1 insures that, up to constant multiples, σ^n is the *only* such eigenvector. The statement about univalence is important since it implies that if φ is univalent, then the eigenfunction σ "linearizes" φ to $M_{\varphi'(p)}$ in accordance with our discussion at the beginning of this section.

To prove Theorem 2.4 it is enough to consider the case where the fixed point p is the origin. Once this is finished, given a Koenigs map whose fixed point p is not 0 we apply the "$p = 0$" result to $\psi = \alpha_p \circ \varphi \circ \alpha_p$, which fixes the origin, and use the automorphism α_p to translate the results for ψ into results for φ. I leave the details as an exercise.

Toward the proof of Theorem 2.4. Before jumping into the proof of this result, it is important to understand the roles played by its various hypotheses. For example, it is desirable that φ have an interior fixed point because of part (b) of Exercise 2.1, which shows that if φ has no such fixed point then C_φ can have eigenvalues of infinite multiplicity. Part (a) of that same exercise shows that it is important for φ *not* to be an automorphism of U, since otherwise there may again be eigenvalues of infinite multiplicity. Here is a preliminary result that gets us started toward further understanding.

2.5. **Lemma.** *Suppose φ is a nonconstant holomorphic self-map of U, and that and $C_\varphi f = \lambda f$ for some $f \in H(U)$ and $\lambda \in \mathbf{C}$. Then:*

(a) $\lambda \neq 0$.

(b) *If f is not constant, and φ fixes $p \in U$, then $\lambda \neq 1$, and $f(p) = 0$.*

Proof. Part (a) is just the statement that C_φ is one-to-one, which we proved in the last section.

For part (b), since φ is not an automorphism we know from Theorem 1.11 and the Exercise that follow it that the iterates φ_n converge to p uniformly on compact subsets of U. Now if λ were equal to 1, then we would have $C_\varphi f = f$, and therefore for each $z \in U$:

$$f(z) = C_\varphi^n f(z) = f(\varphi_n(z)) \to f(p),$$

so $f \equiv f(p)$, which contradicts our hypothesis that f is not constant.

To see that $f(p) = 0$ just set $z = p$ in the eigenvalue equation $f(\varphi(z)) = \lambda f(z)$ to get $f(p) = \lambda f(p)$ which, since $\lambda \neq 1$, gives the desired result. □

We next turn to the question of why the eigenvalues in Theorem 2.4 are what they are.

2.6. **Proposition.** *Suppose φ is a holomorphic self-map of U that fixes a point $p \in U$. Suppose that $C_\varphi f = \lambda f$ for some nonconstant $f \in H(U)$ and some $\lambda \in \mathbf{C}$. Then $\lambda = \varphi'(p)^n$ for some $n = 0, 1, 2, \ldots$.*

Proof. Here is the proof for the special case $p = 0$. I leave it to you to employ the conformal automorphism α_p to translate the result into one for arbitrary $p \in U$.

The eigenfunction f is not constant, so by the previous result it vanishes at the origin, hence there exists a positive integer n such that

$$f(z) = a_n z^n + a_{n+1} z^{n+1} + \cdots \ ,$$

where $a_n \neq 0$. Upon solving the eigenvalue equation $f \circ \varphi = \lambda f$ for λ we obtain for each $z \in U$:

$$\lambda = \frac{f(\varphi(z))}{f(z)} = \left(\frac{\varphi(z)}{z}\right)^n \frac{a_n + a_{n+1}\varphi(z) + a_{n+2}\varphi(z)^2 + \cdots}{a_n + a_{n+1}z + a_{n+2}z^2 + \cdots}.$$

Clearly the right-hand side of this equation converges to $\varphi'(0)^n$ as $z \to 0$, which establishes the desired result. □

Note that in the last result we did not rule out the possibility that $\varphi'(p) = 0$. Since 0 cannot be an eigenvalue of C_φ (for φ nonconstant), the Proposition tells us that if $\varphi'(p) = 0$ then 1 is the only eigenvalue of C_φ. Or, to put it differently, if C_φ has a nonconstant eigenfunction, then $\varphi'(p) \neq 0$.

So far we have justified all the elements that go into the definition of "Koenigs map". Now let's tackle the question of eigenvalue multiplicity.

2.7. **Proposition.** *If φ is a Koenigs map and λ an eigenvalue of C_φ, then λ has multiplicity one, in other words: if $g \in H(U)$ and $C_\varphi g = \lambda g$, then g is a constant multiple of f.*

Proof. Note that we have already established a special case of this: in the proof of Lemma 2.5 we saw that the only eigenvectors for the eigenvalue 1 are the constant functions.

To make further progress, successively differentiate both sides of Schröder's equation, and evaluate each result at $z = 0$. The calculation shows that for every $n \geq 2$, the quantity $(\lambda - \lambda^n) f^{(n)}(0)$ is given by an expression that involves the derivatives $\varphi^{(k)}(0)$ for $1 \leq k \leq n$, and $f^{(j)}(0)$ for $1 \leq j \leq n-1$. Since λ is neither 0 nor 1, an induction argument shows that for $n \geq 2$ the derivative $f^{(n)}(0)$ is determined solely by φ and $f'(0)$. So given φ, the coefficients of f in its Taylor expansion about the origin are determined solely by $f'(0)$ (recall that, because $f(0) = 0$, the constant coefficient is zero). This completes the proof. □

2.8. **Exercise.** Go back to Exercises 2.2 and 2.3. For the first of these, the work we have done so far shows that the eigenvalues $\{\lambda^n\}_0^\infty$ found there are the *only* eigenvalues for C_{φ_λ}, and they have multiplicity one. Show that by contrast, the mappings created in Exercise 2.3 induce composition operators with many more eigenvalues, all of infinite multiplicity! (Hint: go over to the right half-plane Π and consider the composition operator induced by M_λ on $H(\Pi)$.)

Now we understand everything about what eigenvalues for C_φ have to look like when φ is a Koenigs function. What we have *not yet* proved is that there *are* any nontrivial (i.e. not $= 1$) eigenvalues, or, equivalently, nontrivial (not $\equiv$ constant) eigenvectors. We remedy this deficiency with the next result, which almost completes our proof of the Theorem of Koenigs.

2.9. **Theorem.** *Suppose φ is a Koenigs map with fixed point $p \in U$. Then there exists $\sigma \in H(U)$ such that $\sigma \circ \varphi = \varphi'(p)\sigma$.*

Proof. As usual, I'll do the proof only for the case $p = 0$. The function σ will be obtained as the limit of a sequence of "normalized iterates" of φ. As previously discussed, we may take $p = 0$, so that $\varphi(0) = 0$. By the Schwarz Lemma, $0 < |\varphi'(0)| < 1$. Let $\lambda = \varphi'(0)$, and for each positive integer n set $\sigma_n = \lambda^{-n}\varphi_n$ (noting that while the subscript n on the right-hand side of the equation denotes an iterate, the one on the left does not). Since $\sigma_n \circ \varphi = \lambda\sigma_{n+1}$ for each n, the solution σ we seek will be the limit of the sequence of functions $\{\sigma_n\}$, if only we can prove that this limit exists uniformly on compact subsets of U.

To do so, write

$$\sigma_n(z) = z \cdot \frac{\varphi(z)}{\lambda z} \cdot \frac{\varphi_2(z)}{\lambda\varphi(z)} \cdot \;\cdots\; \cdot \frac{\varphi_n(z)}{\lambda\varphi_{n-1}(z)} = z \prod_{j=0}^{n-1} F(\varphi_j(z)),$$

where in the product on the right, $\varphi_0(z) \equiv z$, and $F(z) = \varphi(z)/\lambda z$. Thus our task is to prove that the infinite product

$$\prod_{j=0}^{\infty} F(\varphi_j(z))$$

converges uniformly on compact subsets of U. For this it is enough to show that the infinite series

$$\sum_{j=0}^{\infty} |1 - F(\varphi_j(z))| \tag{2.3}$$

converges likewise. We estimate the size of each term. Since $\varphi(0) = 0$, the function F is holomorphic on U. Let $\|F\|_\infty$ denote the supremum of the values of $|F|$ over the unit disc. By the Maximum Principle, $\|F\|_\infty \le |\lambda^{-1}|$, hence

$$\|1 - F\|_\infty \le 1 + \|F\|_\infty \le 1 + \frac{1}{|\lambda|} \stackrel{\text{def}}{=} A.$$

Finally, the definition $\lambda = \varphi'(0)$ forces $F(0) = 1$, so the Schwarz Lemma, applied to the function $(1 - F)/A$, yields

$$|1 - F(z)| \le A|z| \qquad (z \in U). \tag{2.4}$$

Now fix $0 < r < 1$. In the proof of Theorem 1.11 we observed that a slight refinement of the Schwarz Lemma produces a constant $\delta < 1$ such that for each non-negative integer j,

$$|\varphi_j(z)| \le \delta^j |z|$$

whenever $|z| \le r$. Upon substituting this inequality in (2.4) we see that

$$|1 - F(\varphi_j(z))| \le A|\varphi_j(z)| \le A\delta^j|z| \qquad (z \in r\overline{U}).$$

So on the closed disc $r\overline{U}$, each term of the series (2.3) is bounded uniformly by the corresponding term of a convergent geometric series, hence the original series converges uniformly on that disc. This establishes the desired convergence for the sequence $\{\sigma_n\}$. □

We refer to the function σ constructed above as the *principal eigenfunction* of C_φ. It is the (essentially unique) eigenvector corresponding to the largest nontrivial eigenvalue, namely $\varphi'(0)$.

All that's left in our quest to prove the Theorem of Koenigs is the result about univalence.

2.10. **Corollary.** *If φ is a* univalent *Koenigs map, then the principle eigenfunction of C_φ is also univalent*

Proof. The univalence of φ gets inherited by each iterate φ_n, and therefore by each normalized iterate $\sigma_n = \varphi_n/\lambda^n$. Now suppose the limit function σ were *not* univalent. Then there would be points $a, b \in U$ with $\sigma(a) = \sigma(b)$. Call this common value w. Fix a number r with $\max\{|a|, |b|\} < r < 1$, so that φ does not take the value w on the circle $\gamma = \{|z| = r\}$. Give γ the positive orientation. By the Argument Principle, the integral

$$I \stackrel{\text{def}}{=} \frac{1}{2\pi i} \int_\gamma \frac{\sigma'(z)}{\sigma(z) - w}\, dz$$

counts the number of points in the disc $\{|z| < r\}$ that φ maps onto w, so this integral is at least 2. On the other hand, the same reasoning applied to the univalent maps σ_n shows that for each n,

$$I_n \stackrel{\text{def}}{=} \frac{1}{2\pi i} \int_\gamma \frac{\sigma_n'(z)}{\sigma_n(z) - w}\, dz \leq 1.$$

Since $\sigma_n \to \sigma$ uniformly on compact subsets of U, the integrand of I_n converges uniformly on γ to the integrand of I, and hence $I_n \to I$. But preceding estimates show that this cannot happen, so the assumption that σ is not univalent has led to a contradiction. □

The Theorem of Koenigs is now proved.

3. Composition operators on H^2

The scene now shifts to the *Hardy space* H^2, a subspace of $H(U)$ that is a Hilbert space. This is arguably the best place to study the interaction between the theory of linear operators and analytic function theory. The main result of this section is that every composition operator restricts to a continuous mapping of H^2 into itself. At the core of this result lies the famous *Littlewood Subordination Principle.*

3.1. **Taylor series.** For $f \in H(U)$ and every non-negative integer n, let $\hat{f}(n) = f^{(n)}(0)/n!$. Then the series $\sum_{n=0}^{\infty} \hat{f}(n) z^n$ is the Taylor series of f with center at the origin: it converges uniformly on compact subsets of U to f.

3.2. **The Hardy space.** The *Hardy space* H^2 is the collection of functions $f \in H(U)$ with $\sum_{n=0}^{\infty} |\hat{f}(n)|^2 < \infty$.

3.3. **Exercise.** For α real let $f_\alpha(z) = (1-z)^{-\alpha}$. Show that $f_\alpha \in H^2$ if and only if $\alpha < 1/2$. (Hint: Use the Binomial theorem and Stirling's formula to show that $\hat{f}_\alpha(n) \approx n^{\alpha-1}$.

We equip H^2 with the norm that is naturally associated with its definition:

$$\|f\| = \left(\sum_{n=0}^{\infty} |\hat{f}(n)|^2 \right)^{1/2}, \tag{3.1}$$

and note that this norm arises from the natural inner product

$$\langle f, g \rangle \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} \hat{f}(n) \overline{\hat{g}(n)} \qquad (f, g \in H^2). \tag{3.2}$$

Let T be the "Taylor transformation" from H^2 into the sequence space ℓ^2 defined by $Tf = \hat{f}$. The mapping T is clearly linear, and from the definition of the H^2 norm, it is an isometry: $\|Tf\| = \|f\|$ for every $f \in H^2$.

3.4. **Proposition.** *T maps H^2 onto ℓ^2. In particular, H^2 is a Hilbert space in the inner product (3.2).*

Proof. Because square-summable sequences are bounded, a simple geometric series estimate shows that if the complex sequence $\vec{a} = \{a_n\}_0^\infty$ lies in ℓ^2, then the associated power series $\sum_{n=0}^{\infty} a_n z^n$ converges uniformly on compact subsets of U to an analytic function f. By the uniqueness of power series representations, $a_n = \hat{f}(n)$ for every n, hence $Tf = \vec{a}$, so $T(H^2) = \ell^2$. □

Thus H^2 is the sequence space ℓ^2, disguised as a space of analytic functions. Note in particular that:

3.5. **Proposition.** *The sequence of monomials $\{z^n : n = 0, 1, 2, \dots\}$ is an orthonormal basis for H^2.*

Some properties of the functions in H^2 can be easily discerned from the definition of the space. Here is one.

3.6. **Growth Estimate.** *For every $f \in H^2$ and $z \in U$,*

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

Proof. Use successively the triangle inequality and the Cauchy-Schwarz Inequality on the power series representation for f:

$$\begin{aligned} |f(z)| &= |\sum_{n=0}^{\infty} \hat{f}(n) z^n| \\ &\leq \sum_{n=0}^{\infty} |\hat{f}(n)||z|^n \\ &\leq \left(\sum_{n=0}^{\infty} |\hat{f}(n)|^2\right)^{1/2} \left(\sum_{n=0}^{\infty} |z|^2\right)^{1/2} \\ &= \|f\| \frac{1}{(1-|z|^{2n})^{1/2}}. \end{aligned}$$

□

3.7. **Corollary.** *Convergence in H^2 implies uniform convergence on compact subsets of U.*

Proof. Suppose $\{f_n\}$ is a sequence of functions in H^2, f is a function in H^2, and $\|f_n - f\| \to 0$. Our goal is to show that $f_n \to f$ uniformly on compact subsets of U.

For this, suppose K is a compact subset of U. Let $r = \max\{|z| : z \in K\}$. Then for $z \in K$, the Growth Estimate yields:

$$|f_n(z) - f(z)| \leq \frac{\|f_n - f\|}{(1-|z|^2)^{1/2}} \leq \frac{\|f_n - f\|}{(1-|r|^2)^{1/2}},$$

which shows that as $n \to \infty$,

$$\max_{z \in K} |f_n(z) - f(z)| \leq \frac{\|f_n - f\|}{(1-|r|^2)^{1/2}} \to 0,$$

i.e. that $f_n \to f$ uniformly on K. □

3.8. **Exercise.** For $\alpha \in \mathbf{C}$, let $f_\alpha(z) = (1-z)^\alpha$. Show that $f_\alpha \notin H^2$ whenever $\operatorname{Re}\alpha \leq -1/2$.

However some properties of H^2 do not follow easily from the definition. For example, is every bounded analytic function in H^2? In order to answer this question reasonably, we need a different description of the norm.

3.9. **Proposition.** *A function $f \in H(U)$ belongs to H^2 if and only if*

$$\lim_{r \to 1-} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 \, d\theta < \infty.$$

When this happens, the limit of integrals on the left is $\|f\|^2$.

Proof. The functions $e^{in\theta}$ form an orthonormal set in the space $L^2([0, 2\pi])$, hence for each $0 \leq r < 1$ the integral on the right is $\sum_{n=0}^{\infty} |\hat{f}(n)|r^{2n}$. The result now follows from the monotone convergence theorem. □

3.10. **Exercise.** Show that the function f_α of Exercise 3.8 is in H^2 whenever $\operatorname{Re} \alpha > -1/2$.

It is now an easy matter to show that every bounded function in $H(U)$ belongs to H^2. In fact, we can do better. Let H^∞ denote the collection of bounded analytic functions on U, and for $b \in H^\infty$ let $\|b\|_\infty = \sup_{z \in U} |b(z)|$. The integral representation given above for the H^2 norm shows immediately:

3.11. **Proposition.** *If $b \in H^\infty$ and $f \in H^2$ then $bf \in H^2$ and $\|bf\| \leq \|b\|_\infty \|f\|$.*

In particular, upon taking $f \equiv 1$ we obtain:

3.12. **Corollary.** *If $b \in H^\infty$ then $b \in H^2$ with $\|b\| \leq \|b\|_\infty$.*

3.13. **Multiplication operators act on H^2.** Proposition 3.11 reveals an interesting class of linear transformations on H^2. For $b \in H^\infty$ let M_b denote the operator of (pointwise) multiplication by b. That is, $M_b f = bf$. Clearly M_b, when viewed as a mapping on all of $H(U)$, is linear (note that for this we don't need b to be bounded). According to Proposition 3.11 M_b maps H^2 into itself, with $\|M_b f\| \leq \|b\|_\infty \|f\|$ for each $f \in H^2$. It is not difficult to show from this inequality that that M_b is continuous on H^2 (see §3.18 and Proposition 3.24 below). We call M_b the *multiplication operator* induced by b. The most famous of these is the one induced by the identity map $b(z) \equiv z$. Because it shifts the Taylor coefficients of any function on which it acts one unit to the right, this operator of "multiplication by z" is often called the "Forward Shift."

3.14. **Do Composition operators act on H^2?** This is not a trivial question. Suppose you have $f \in H^2$ and want to determine if $C_\varphi f \in H^2$. Using the definition of H^2 we would substitute $\varphi(z)$ for z in the power series expansion of f, expand the various powers of the power series of φ by the binomial theorem, and regroup the resulting double series to identify the new powers of z, which are now complicated numerical series involving the coefficients of f and those of the powers of φ. Done this way, there seems to be no reason why $C_\varphi f$ should be in H^2. A calculation using the alternate characterization of H^2 provided by Proposition 3.9 fares just as badly, since it raises the specter of an unpleasant, and possibly non-univalent, change of variable in an integral.

After these pessimistic observations, it is remarkable that composition operators *do* preserve the space H^2, and do so continuously. The key to this is the following result, proved by Littlewood and published in 1925.

3.15. **Littlewood's Subordination Theorem.** *Suppose φ is a holomorphic self-map of U and $\varphi(0)=0$. Then $C_\varphi f \in H^2$, and $\|C_\varphi f\| \leq \|f\|$ for every $f \in H^2$.*

Proof. The proof is helped significantly by the *backward shift operator* B, defined on H^2 by

$$Bf(z) = \sum_{n=0}^{\infty} \hat{f}(n+1)z^n \qquad (f \in H^2).$$

The name comes from the fact that B shifts the power series coefficients of f one unit to the left, and drops off the constant term. Clearly, $\|Bf\| \leq \|f\|$ for each $f \in H^2$, and one might expect this fact to play an important role in the proof, but surprisingly it does not! Only the following two identities are needed, and they hold for any $f \in H(U)$:

$$f(z) = f(0) + zBf(z) \qquad (z \in U), \tag{3.3}$$

$$B^n f(0) = \hat{f}(n) \qquad (n = 0, 1, 2, \dots). \tag{3.4}$$

To begin the proof, suppose first that f is a (holomorphic) polynomial. Then $f \circ \varphi$ is bounded on U, so by the work of the last section there is no doubt that it lies in H^2; the real issue is its *norm*.

We begin the norm estimate by substituting $\varphi(z)$ for z in (3.3) to obtain

$$f(\varphi(z)) = f(0) + \varphi(z)(Bf)(\varphi(z)) \qquad (z \in U).$$

Let us rewrite this equation in the language of composition and multiplication operators:

$$C_\varphi f = f(0) + M_\varphi C_\varphi B f\,. \tag{3.5}$$

At this point, the assumption $\varphi(0) = 0$ makes its first (and only) appearance. It asserts that all the terms of the power series for φ have a common factor of z, hence the same is true for the second term on the right side of equation (3.5), rendering it orthogonal in H^2 to the constant function $f(0)$. Thus,

$$\|C_\varphi f\|^2 = |f(0)|^2 + \|M_\varphi C_\varphi B f\|^2 \leq |f(0)|^2 + \|C_\varphi Bf)\|^2, \tag{3.6}$$

where the last inequality follows from Proposition 3.11 above (since $\|\varphi\|_\infty \leq 1$. Now successively substitute $Bf, B^2 f, \cdots$ for f in (3.6) to obtain:

$$\begin{array}{rcl} \|C_\varphi Bf\|^2 & \leq & |Bf(0)|^2 + \|C_\varphi B^2 f\|^2 \\ \|C_\varphi B^2 f\|^2 & \leq & |B^2 f(0)|^2 + \|C_\varphi B^3 f\|^2 \\ & \vdots & \qquad\qquad \vdots \\ \|C_\varphi B^n f\|^2 & \leq & |B^n f(0)|^2 + \|C_\varphi B^{n+1} f\|^2. \end{array}$$

Putting all these inequalities together, we get

$$\|C_\varphi f\|^2 \leq \sum_{k=0}^{n} |(B^k f)(0)|^2 + \|C_\varphi B^{n+1} f\|^2$$

for each non-negative integer n.

Now recall that f is a polynomial. If we choose n be the degree of f, then $B^{n+1}f = 0$, and this reduces the last inequality to

$$\begin{aligned} \|C_\varphi f\|^2 &\leq \sum_{k=0}^{n} |(B^k f)(0)|^2 \\ &= \sum_{k=0}^{n} |\hat{f}(k)|^2 \\ &= \|f\|^2, \end{aligned}$$

where the middle line comes from property (3.4) of the backward shift. This shows that C_φ is an H^2-norm contraction, at least on the vector space of holomorphic polynomials.

To finish the proof, suppose $f \in H^2$ is not a polynomial. Let $f_n(z) = \sum_{k=0}^{n} \hat{f}(k) z^k$, the n-th partial sum of the Taylor series of f. Then $f_n \to f$ in the norm of H^2, so by Corollary 3.7 $f_n \to f$ uniformly on compact subsets of U, hence $f_n \circ \varphi \to f \circ \varphi$ in the same manner. It is clear that $\|f_n\| \leq \|f\|$, and we have just shown that $\|f_n \circ \varphi\| \leq \|f_n\|$. Thus for each fixed $0 < r < 1$ we have

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} |f_n(\varphi(re^{i\theta}))|^2 \, d\theta &= \lim_{n\to\infty} \frac{1}{2\pi}\int_0^{2\pi} |f_n(\varphi(re^{i\theta}))|^2 \, d\theta \\ &\leq \limsup_{n\to\infty} \|f_n \circ \varphi\| \\ &\leq \limsup_{n\to\infty} \|f_n\| \\ &\leq \|f\|. \end{aligned}$$

To complete the proof, let r tend to 1, and appeal one last time to Proposition 3.9. □

To prove that C_φ is bounded even when φ does not fix the origin, we utilize the conformal automorphisms α_p introduced in Definition 1.1 to move points of U from where they are to where we want them. For each point $p \in U$, recall that α_p takes U onto itself, interchanges p with the origin, and is its own inverse. Write $p = \varphi(0)$. Then the holomorphic function $\psi = \alpha_p \circ \varphi$ takes U into itself and fixes the origin. By the self-inverse property of α_p we have $\varphi = \alpha_p \circ \psi$, and this translates into the operator equation $C_\varphi = C_\psi C_{\alpha_p}$. We have just seen that C_ψ maps H^2 into itself. Thus, the fact that C_φ does the same will follow from the first sentence of the next result.

3.16. **Lemma.** *For each $p \in U$ the operator C_{α_p} maps H^2 into itself. Moreover,*

$$\|C_{\alpha_p}\| \leq \left(\frac{1+|p|}{1-|p|}\right)^{\frac{1}{2}}.$$

Proof. Suppose first that f is holomorphic in a neighborhood of the closed unit disc, say in $RU = \{|z| < R\}$ for some $R > 1$. Then the limit in formula (3.9) can be passed inside the integral sign, with the result that

$$\|f\|^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(e^{i\theta})|^2 d\theta.$$

This opens the door to a simple change of variable in which the self-inverse property of α_p figures prominently:

$$\begin{aligned}
\|f\circ\alpha_p\|^2 &= \frac{1}{2\pi}\int_0^{2\pi} |f(\alpha_p(e^{i\theta}))|^2 d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} |f(e^{it})|^2 |\alpha_p'(e^{it})| dt \\
&= \frac{1}{2\pi}\int_0^{2\pi} |f(e^{it})|^2 \, \frac{1-|p|^2}{|1-\bar{p}e^{it}|^2} dt \\
&\leq \frac{1-|p|^2}{(1-|p|)^2}\cdot\left(\frac{1}{2\pi}\int_0^{2\pi} |f(e^{it})|^2 dt\right) \\
&= \frac{1+|p|}{1-|p|}\cdot\|f\|^2.
\end{aligned}$$

Thus the desired inequality holds for all functions holomorphic in RU; in particular it holds for polynomials. It remains only to transfer the result to the rest of H^2, and for this we simply repeat the argument used to finish the proof of Littlewood's Subordination Theorem. □

At this point we have assembled everything we need to show that composition operators map H^2 into itself.

3.17. **Theorem.** *Suppose φ is a holomorphic self-map of U. Then C_φ is a bounded operator on H^2, and*

$$\|C_\varphi f\| \leq \sqrt{\frac{1+|\varphi(0)|}{1-|\varphi(0)|}}\|f\|.$$

for every $f \in H^2$.

Proof. As outlined earlier, we have $C_\varphi = C_\psi C_{\alpha_p}$, where $p = \varphi(0)$, and ψ fixes the origin. Since each of the operators on the right-hand side of this equation sends H^2 into itself, the same is true of C_φ.

As for the inequality, this follows from Lemmas 3.15 and 3.16. I leave the details to you.

□

3.18. **Normed spaces and bounded linear transformations.** Suppose that X is a *normed space,* that is, a vector space over the real or complex field on which there is defined a functional $\|\cdot\| : X \to [0,\infty)$ with these properties:

- $\|ax\| = |a|\,\|x\|$ for every $x \in X$ and every scalar a.
- $\|x+y\| \le \|x\| + \|y\|$ for every pair of vectors $x, y \in X$.
- $\|x\| = 0 \implies x = 0$.

Recall that every norm determines a translation-invariant *metric* on the underlying space: $d(x,y) = \|x-y\|$. If this metric is complete, we call the underlying space (with its norm) a *Banach space.* Every Hilbert space is a Banach space. Other examples that you have probably seen are: the sequence spaces ℓ^p for $1 \le p <\le \infty$ and the space $C(K)$, where K is a compact metric space and the norm is the "max-norm".

3.19. **Exercise.** Show that the space H^∞ of bounded analytic functions on the unit disc is a Banach space in the "sup norm": $\|f\|_\infty = \sup_{z\in U} |f(z)|$.

If X and Y are normed spaces and $T : X \to Y$ is a linear transformation, then T is said to be *bounded* if there is a non-negative constant C such that

$$\|Tx\| \le C\|x\| \quad \text{for every } x \in X \tag{3.7}$$

Note that in order to conserve notation we use the same symbol for the norm in X (on the right side of the equation above) that we use for the norm in Y (on the left side). The infimum of all the numbers C that work in the above definition is called the *norm* of T.

3.20. **Exercise.** Show that $\|T\|$ is also a value of C that works in (3.7) above. In other words, if T is a bounded linear transformation between normed spaces, then $\|Tx\| \le \|T\|\,\|x\|$.

Note that the term "bounded" for a linear transformation does not mean that $\|Tx\|$ is bounded as x runs over the whole space. It is an easy exercise to show that the only zero-transformation has this extreme form of boundedness.

3.21. **Exercise.** Show that a linear transformation between normed spaces is bounded if and only if the image of every bounded set is bounded(by a *bounded subset* of a normed space we mean a set A for which $\sup_{x\in A} \|x\| < \infty$).

3.22. **Exercise.** Show that if T is a bounded linear transformation, then

$$\|T\| = \sup\{\|Tx\| : \|x\| \le 1\},$$

and that the inequality on the right-hand side of (3.7) can be replaced by both equality and strict inequality.

3.23. **Exercise.** Show that every linear transformation on a finite dimensional Hilbert space is bounded.

Examples of bounded linear transformations that we have seen so far are:

(a) The "Taylor transform" $T : H^2 \to \ell^2$ defined just before the statement of Proposition 3.4). Since T is an isometry, $\|T\| = 1$. Note, however, that non-isometries can also have norm 1. The backward shift operator B that showed up in the proof of Littlewood's Theorem (§3.15) is a case in point (exercise).

(b) Multiplication operators on H^2. The work of §3.13 actually showed that if $b \in H^\infty$ then M_b is a bounded bounded operator on H^2, and $\|M_b\| \leq \|b\|_\infty$. In fact there is equality here (exercise).

(c) Theorem 3.17 showed that for every holomorphic self-map φ of U, the composition operator C_φ is bounded on H^2, with

$$\|C_\varphi\| \leq \left(\frac{1+|\varphi(0)|}{1-|\varphi(0)|}\right)^{1/2}.$$

There is no known formula for the norm of a general composition operator.

The next result shows that all the above-mentioned operators are continuous.

3.24. **Proposition.** *Every bounded linear transformation between normed spaces is continuous.*

Proof. Suppose $T : X \to Y$ is bounded, so it satisfied (3.7) above. Then for any pair of vectors $x, y \in X$,

$$\|Tx - Ty\| = \|T(x-y)\| \leq C\|x-y\|,$$

which shows that T is continuous (in fact, uniformly continuous). □

In fact, boundedness for a linear transformation between normed spaces is *equivalent* to continuity.

3.25. **Theorem.** *Every continuous linear transformation between normed spaces is bounded.*

Proof. Suppose $T : X \to Y$ is continuous. Then the inverse image of the unit ball in Y contains an open ball $\{x \in X : \|x - x_0\| < \varepsilon\}$, for some $x_0 \in X$ and $\varepsilon > 0$. Thus

$$\|x - x_0\| < \varepsilon \implies \|Tx\| < 1.$$

Upon setting $z = (x - x_0)/\varepsilon$, so that $x = \varepsilon z + x_0$, we see from the implication above that whenever $\|z\| < 1$:

$$\begin{aligned} 1 &> \|T(\varepsilon z + x_0)\| \\ &= \|\varepsilon T(z) + T(x_0)\| \\ &\geq \varepsilon\|T(z)\| - \|T(x_0)\|, \end{aligned}$$

hence

$$\|T(z)\| \le \varepsilon^{-1}(1 + \|T(x_0)\|).$$

For arbitrary $z \in X$ ($z \neq 0$) we apply the last inequality to the unit vector $z/\|z\|$ to get $\|T(z)\| \le C\|z\|$, where $C = \varepsilon^{-1}(1 + \|T(x_0)\|)$. Thus T is a bounded linear transformation. □

This section closes with some further exercises on the boundedness of composition operators.

3.26. **Exercise.** Let $\ell^1(U)$ denote the space of absolutely summable complex sequences, but now regarded as a space of analytic functions. That is,

$$\ell^1(U) = \{f \in H(U) : \|f\|_1 \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} |\hat{f}(n)| < \infty\}$$

Show that the map $Tf = \hat{f}$ is a linear isometry of $\ell^1(U)$ onto the sequence space ℓ^1, and that the following version of Littlewood's Subordination Theorem holds for $\ell^1(U)$:

> *Suppose $\varphi \in \ell^1(U)$ and $\|\varphi\|_1 \le 1$. Then φ is a holomorphic self-map of U, and C_φ is a contraction on $\ell^1(U)$.*

3.27. **Exercise.** Show that for every $0 < \alpha < 1/2$, the function $\left(\frac{1+z}{1-z}\right)^\alpha$ belongs to H^2.

3.28. **Exercise.** Suppose f and g are holomorphic on U, with g univalent and $f(U) \subset g(U)$. Show that if g belongs to H^2, then so does f. (Hint: first show that $g = f \circ \varphi$ for some holomorphic self-map φ of U with $\varphi(0) = 0$).

3.29. **Exercise.** Use the results of the previous two exercises to show that if f is holomorphic on U, and $f(U)$ is contained in an angular sector with vertex angle less than $\pi/2$ radians, then $f \in H^2$.

4. Introduction to the Compactness Problem

Recall that a linear transformation between normed spaces is continuous if and only if it is bounded (Proposition 3.24 and Theorem 3.25), i.e., if and only if the image of every bounded set is bounded (Exercise 3.21). The most important subclass of bounded operators are the *compact* ones.

4.1. **Definition.** A linear transformation between normed spaces is said to be *compact* if the image of every bounded set is relatively compact.

To say that a set is "relatively compact" means that it has compact closure. In the language of sequences this says that every sequence taken from that set has a convergent subsequence.

Since compact sets are bounded, it follows immediately that compact operators are bounded.

4.2. **Exercise.** Show that a linear transformation is compact if and only if the image of the unit ball, or more generally, if the image of *some* ball is relatively compact.

We have already noted that every linear operator on a finite dimensional Hilbert space is bounded (Exercise 3.23). Because of the Bolzano-Weierstrass theorem, every such operator is therefore compact. In fact, in this argument only the range space is important:

> *If a bounded linear operator on a Hilbert space has finite dimensional range, then the operator is compact.*

On the other hand,

4.3. **Exercise.** The identity operator on an infinite dimensional Hilbert space is *not* compact. (Suggestion: Use the existence of an orthonormal basis to show that the closed unit ball is not compact).

The connection between compact operators and finite dimensional ones is, in fact, quite strong. One of the most important results about compact operators is this:

4.4. **Theorem.** *If X is a Banach space and $T : X \to X$ is compact, then for each nonzero $\lambda \in \mathbf{C}$, the following are equivalent:*

(a) $T - \lambda I$ *is invertible.*
(b) $T - \lambda I$ *is one to one.*
(c) $T - \lambda I$ *maps X onto itself.*

This result, which we won't prove here, asserts that perturbations of the identity by compact operators have the same invertibility characteristics as operators on a finite dimensional space.

Once you know that a collection of operators between normed spaces are all bounded, it makes sense to try to determine which of these is compact. This is a particularly interesting problem for composition operators on H^2. Indeed, "C_φ compact on H^2 means that the operator squeezes the unit ball of H^2 into a relatively compact (=

"small") subset. The question is: how much does φ have to squeeze the unit disc into itself in order for this to happen?

Here are two amusing examples to get us started: If $\varphi(z) = z$ then C_φ is the identity, which, according to Exercise 4.3, is not compact. On the other hand, if $\varphi(z) \equiv$ constant then C_φ has one dimensional range (the constants), and is compact. To examine what lies between we need a "sequential" characterization of compactness that was first observed by H.J. Schwartz (no relation to the Schwarz Lemma) in 1969.

4.5. **Theorem.** *C_φ is compact on H^2 if and only: whenever $\{f_n\}$ is a bounded sequence in H^2 that converges to zero uniformly on compact subsets of U, then $\|C_\varphi f_n\| \to 0$.*

The proof of this result depends on some deep results about uniform convergence. All of this will be developed shortly, but right now perhaps a better idea would be to give some applications. The first one generalizes the fact that $\varphi(z) \equiv z$ induces a noncompact operator.

4.6. **Proposition.** *Suppose C_φ is compact Then the set of points $e^{i\theta} \in \partial U$ at which $\lim_{r\to 1-} |\varphi(e^{i\theta})| = 1$ has Lebesgue measure zero.*

Proof. Suppose that $E_\varphi = \{e^{i\theta} : \lim_{r\to 1-} |\varphi(e^{i\theta})| = 1\}$ has positive measure. We will show that C_φ is not compact.

Note first that, since $|\varphi| < 1$ on U, the "limit" in the definition of E_φ may be replaced by "liminf". Now the monomial sequence $\{z^n\}_0^\infty$ obeys the hypothesis of Theorem 4.5, so by our hypothesis on C_φ, the image sequence, which is $\{\varphi^\}$ converges to zero in H^2. Since φ is continuous on $\overline{U}$, as $r \to 1-$ we have $\varphi(re^{i\theta}) \to \varphi(e^{i\theta})$ uniformly over the unit circle, and the same is true for φ^n. Thus the integral representation of the H^2 norm provided by Theorem 3.9, along with Fatou's Lemma, yields

$$
\begin{aligned}
\|\varphi^n\| &= \lim_{r\to 1-} \frac{1}{2\pi} \int_0^{2\pi} |\varphi(re^{i\theta})|^{2n}\, d\theta \\
&= \frac{1}{2\pi} \int_0^{2\pi} |\liminf_{r\to 1-} \varphi(e^{i\theta})|^{2n}\, d\theta \\
&= \frac{1}{2\pi} \int_0^{2\pi} |(\liminf_{r\to 1-} \varphi(e^{i\theta})|)^{2n}\, d\theta \\
&\geq \frac{1}{2\pi} \int_{E_\varphi} |\varphi(e^{i\theta})|^{2n}\, d\theta \\
&= \frac{1}{2\pi} \text{meas}\,(E_\varphi) > 0.
\end{aligned}
$$

This calculation shows that $\|C_\varphi(z^n)\|$ does not converge to zero, so by Theorem 4.5, C_φ is not compact. ☐

It is known that for every bounded analytic function, and more generally for any H^2 function, the *radial limit*

$$f(e^{i\theta}) \stackrel{\text{def}}{=} \lim_{r\to 1-} f(re^{i\theta})$$

exists for almost every $e^{i\theta}$ in the unit circle. In particular, the limit that shows up in the definition of E_φ in the proof above exists for a.e. point of the unit circle.

4.7. **Exercise.** Construct a holomorphic self-map φ of U, not the identity map, that induces a noncompact composition operator on H^2. (Suggestion: map the unit disc onto the upper half-plane and take a square root.)

The next application of Theorem 4.5 generalizes the fact that constant φ's induce compact composition operators. Recall the notation $\|\varphi\|_\infty = \sup_{z\in U} |f(z)|$.

4.8. **Proposition.** *If $\|\varphi\|_\infty < 1$ then C_φ is compact on H^2.*

Proof. Suppose $\{f_n\}$ is a bounded sequence in H^2 that converges to zero uniformly on compact subsets of U. By Theorem 4.5 it is enough to show that $\|f_n \circ \varphi\| \to 0$. But even more is true: since $\varphi(U)$ is a relatively compact subset of U, $f_n \to 0$ uniformly on $\varphi(U)$, hence

$$\|f_n \circ \varphi\| \le \|f_n \circ \varphi\|_\infty \le \sup_{w\in\varphi(U)} |f_n(w)| \to 0,$$

as desired. □

Observe that we did not need the boundedness of the sequence $\{f_n\}$ for this argument. This suggests that the sufficient condition $\|\varphi\|_\infty < 1$ for compactness of C_φ might not be necessary. We'll see later on that this is indeed the case: there exist maps φ with $\|\varphi\|_\infty = 1$ for which C_φ is compact on H^2.

Before approaching the proof of Theorem 4.5 I present one final application that shows that the necessary condition of Proposition 4.6 does not provide a characterization of compact composition operators on H^2.

4.9. **Proposition.** *Let $\varphi(z) = \frac{1+z}{2}$. Then C_φ is not compact on H^2.*

Proof. Proof. Consider for $0\alpha < 1/2$ the functions $f_\alpha(z) = (1-z)^{-\alpha}$. By Exercise 3.3 they all belong to H^2, and another exercise shows that $\|f_\alpha\| \to \infty$ as $\alpha \to 1/2$. Let $g_\alpha = f_\alpha/\|f_\alpha\|$. Then $\{g_\alpha\}$ is a collection of unit vectors that converges to zero uniformly on compact subsets of U as $\alpha \to 1/2$.

Now each f_α is an eigenvector of C_φ, in fact $C_\varphi f_\alpha = 2^\alpha f_\alpha$. Thus the same is true of g_α, hence as $\alpha \to 1/2$,

$$\|C_\varphi g_\alpha\| = 2^\alpha \to \sqrt{2} \ne 0,$$

so by Theorem 4.5, C_φ is not compact. □

4.10. **Proof of Theorem 4.5: necessity.** This follows quickly from Corollary 3.7. For suppose C_φ is compact on H^2, and that $\{f_n\} \subset H^2$ is bounded and converges to zero uniformly on compact subsets of U. Then the image sequence $\{f_n \circ \varphi\}$ is relatively compact, so it has a subsequence $\{g_k = f_{n_k} \circ \varphi\}$ that converges in the H^2 norm to some function $g \in H^2$. Our goal is to show that $g = 0$. Since H^2 convergence implies uniform convergence on compact subsets of U (Corollary 3.7) we know that $g_k \to g$ in that manner. But $f_{n_k} \to 0$ uniformly in compact subsets of U, hence the same is true of $f_{n_k} \circ \varphi = g_k$, hence $g = 0$, as desired. □

The converse is more difficult. To prove it we need two famous theorems of classical analysis, the first which revolves around the notion of *equicontinuity.* If A is a collection of functions that are continuous on a set $S \subset \mathbf{C}$ (or on any metric space, for that matter), we say S is *equicontinuous* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $x, 4$ is a pair of points in S with $|x - y| < \delta$, the $|f(x) - f(y)| < \varepsilon$ for every $f \in A$. Thus each $f \in A$ is continuous on S, and for each ε the constant δ can be chosen to work for all the functions in A.

4.11. **The Arzela–Ascoli Theorem.** *If K is a compact subset of the plane and $A \subset C(K)$ is uniformly bounded and equicontinuous, then every sequence in A has a uniformly convergent subsequence.*

Proof. Since K is compact, for each positive integer n there is a finite collection $\mathcal{F}_n$ of open discs of radius $1/n$, with centers in K, that covers K. Let S denote the collection of all the centers of all the discs in all the open covers $\mathcal{F}_n$. This is a countable dense subset of K.

Now consider a sequence $\{f_n\}$ in A. Because of the boundedness assumption on A we know that there exists $M > 0$ such that $|f_n(z)| \le M$ for each $z \in K$ and each n. Since bounded sequences of complex numbers have convergent subsequences, a diagonal argument produces a subsequence $g_k = f_{n_k}$ that converges pointwise on S. I claim that $\{g_k\}$ is uniformly Cauchy on K, i.e. a Cauchy sequence in the Banach space $C(K)$.

For this, let $\varepsilon > 0$ be given. Use equicontinuity to choose a positive integer N such that if $z, w \in K$ and $|z - w| < 1/N$, then $|f(z) - f(w)| < \varepsilon/3$ for each $f \in A$. Next, use the pointwise convergence of $\{g_k\}$ on S to choose n_ε such that whenever k and j are integers $> n_\varepsilon$ and a is the center of a disc in $\mathcal{F}_N$ (there are only finitely many discs here!), we have $|g_k(a) - g_j(a)| < \varepsilon/3$. I claim that if $j, k > N_\varepsilon$ then for every $z \in K$ we have $|g_k(z) - g_j(z)| < \varepsilon$, which will prove the desired result: $\{g_k\}$ is uniformly Cauchy on K.

For this, fix $j, k > N_\varepsilon$, and fix a point $z \in K$. Since $\mathcal{F}_n$ is a cover of K there is a disc $\Delta \in \mathcal{F}_n$ that contains z. Let a be the center of Δ, so that $|z - a| < 1/N$. Then

$$\begin{aligned} |g_k(z) - g_j(z)| &\le |g_k(z) - g_k(a)| + |g_k(a) - g_j(a)| + |g_j(a) - g_j(z)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon, \end{aligned}$$

where in the second line the first and last inequality come from the proximity of z to a, and and the middle one from the choice of n_ε. □

4.12. **Exercise.** Show that the converse of Theorem 4.11 holds also: If every sequence in A has a convergent subsequence, then A is uniformly bounded and equicontinuous. (Suggestion: The hypothesis really asserts that A is relatively compact in $C(K)$. An equivalent formulation of relative compactness thus, for every $\varepsilon > 0$ there is a finite covering of A by balls of radius ε with centers in A. These centers are the key.)

4.13. **Theorem.** *Suppose $A \subset H(U)$ is uniformly bounded on a compact subset K of U. Then A is equicontinuous on K.*

Proof. Choose $0 < r < 1$ so that $K \subset \{|z| < r\}$, and let γ be the circle $\{|z| = r\}$, oriented positively, and let

$$d = \operatorname{dist}(K, \gamma) \stackrel{\text{def}}{=} \inf\{|w - z| : w \in K \text{ and } z \in \gamma\}.$$

Since K and γ are disjoint and compact, $d > 0$.

Since γ is compact we know that

$$M \stackrel{\text{def}}{=} \sup\{|f(z)| : z \in \gamma, f \in A\} < \infty.$$

Fix z_1 and z_2 in K, and fix $f \in A$. The Cauchy integral formula, along with a little bit of algebra shows that

$$f(z_1) - f(z_2) = \frac{z_1 - z_2}{2\pi i} \int_\gamma \frac{f(z)}{(z - z_1)(z - z_2)}\, dz$$

from which it follows quickly that

$$|f(z_1) - f(z_2)| \leq \frac{Mr}{d^2} |z_1 - z_2|.$$

Since the constants M,r, and d do not depend on either $f \in A$ or $z_1, z_2 \in K$, this shows that A is equicontinuous on K. □

4.14. **The Osgood-Stieltjes Theorem.** *If $A \subset H(U)$ is uniformly bounded on every compact subset of U, then every sequence in A has a subsequence that converges uniformly on compact subsets of U.*

Proof. This is a standard diagonal argument. Let Δ_j be the closed disc of radius $1 - \frac{1}{j}$, centered at the origin, so U is the union of all these subdiscs. Suppose $\{f_n\}$ is a sequence of functions in A. By Theorems 4.13 and 4.11 there is a subsequence $\{f_{1,n}\}$ of $\{f_n\}$ that converges uniformly on Δ_1. Similarly there exists a subsequence $\{f_{2,n}\}$ of $\{f_{1,n}\}$ that converges uniformly on Δ_2. Continuing in this manner we obtain an infinite matrix $\{f_{m,n}\}$ of functions where:

- Each row is a subsequence of the previous row.
- The first row is a subsequence of the original sequence $\{f_n\}$.
- The n-th row converges uniformly on Δ_n.

The "diagonal sequence" $\{f_{n,n}\}$ is therefore a subsequence $\{f_n\}$ that converges uniformly on each of the closed discs Δ_j, and therefore uniformly on every compact subset of U. $\square$

4.15. **Proof of Theorem 4.5: sufficiency.** Suppose C_φ takes every sequence that is bounded in H^2 and convergent uniformly on compact subsets of U into a sequence that converges to zero in H^2. Our goal is to show that the image of the unit ball of H^2 is relatively compact.

Fix a sequence $\{f_n\}$ in the closed unit ball $\mathcal{B}$ of H^2. According to the Growth Estimate §3.6, $\mathcal{B}$ is bounded uniformly on compact subsets of U, so by Theorem 4.14 the sequence $\{f_n\}$ has a subsequence $g_k = f_{n_k}$ that converges uniformly on compact subsets of U to a function $g \in \mathcal{B}$. Thus $\{g_k - g\} \leq 2$ for every k, and $g_k \to$ uniformly on compact sets, hence the hypothesis on C_φ implies that $\|g_k \circ \varphi - g \circ \varphi\| \to 0$. This shows that every sequence in $C_\varphi(\mathcal{B})$ has a convergent subsequence, i.e., that $C_\varphi(\mathcal{B})$ is relatively compact. This completes the proof. $\square$