Stephan Ramon Garcia
Javad Mashreghi · William T. Ross

Finite Blaschke Products and Their Connections

Springer

Finite Blaschke Products and Their Connections

Stephan Ramon Garcia • Javad Mashreghi
William T. Ross

Finite Blaschke Products and Their Connections

Springer

Stephan Ramon Garcia
Department of Mathematics
Pomona College
Claremont, CA, USA

Javad Mashreghi
Department of Mathematics and Statistics
Laval University
Québec City, QC, Canada

William T. Ross
Department of Mathematics and Computer Science
University of Richmond
Richmond, VA, USA

ISBN 978-3-319-78246-1 ISBN 978-3-319-78247-8 (eBook)
https://doi.org/10.1007/978-3-319-78247-8

Library of Congress Control Number: 2018937693

Mathematics Subject Classification (2010): 30J10, 30J05, 30A78, 30A76, 30A08, 30C15, 30E10, 47A12, 47N50, 47A30

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To our families:
Gizem; Reyhan, and Altay
Shahzad; Dorsa, Parisa, and Golsa
Fiona

Preface

This is a book about a beautiful subject that begins with the topic of Möbius transformations. Indeed, Möbius transformations

$$z \mapsto \frac{az+b}{cz+d}$$

are studied in complex analysis since their mapping properties demonstrate wonderful connections with geometry. These transformations map extended circles to extended circles, enjoy the symmetry principle, come in several types yielding different behavior depending on their fixed point(s), and, through an identification with 2×2 matrices, make connections to group theory and projective geometry. Finite Blaschke products, the focus of this book, are products of certain types of Möbius transformations, the automorphisms of the open unit disk $\mathbb{D}$, namely

$$z \mapsto \xi \frac{w-z}{1-\overline{w}z},$$

where $|w| < 1$ and $|\xi| = 1$ are fixed. These products have an uncanny way of appearing in many areas of mathematics such as complex analysis, linear algebra, group theory, operator theory, and systems theory. This book covers finite Blaschke products and is designed for advanced undergraduate students, graduate students, and researchers who are familiar with complex analysis but who want to see more of its connections to other fields of mathematics. Much of the material in this book is scattered throughout mathematical history, often only appearing in its original language, and some of it has never seen a modern exposition. We gather up these gems and put them together as a cohesive whole, taking a leisurely pace through the subject and leaving plenty of time for exposition and examples. There are plenty of exercises for the reader who not only wants to appreciate the beauty of the subject but to gain a working knowledge of it as well.

In the early twentieth century, the study of infinite products of the form

$$B(z) = \prod_{k \geqslant 1} \frac{|z_k|}{z_k} \frac{z_k - z}{1 - \overline{z_k} z},$$

in which $z_1, z_2, \ldots$ is a sequence in $\mathbb{D}$, was initiated in 1915 by Wilhelm Blaschke (1885–1962). This product converges uniformly on compact subsets of $\mathbb{D}$ if and only if the zero sequence z_k satisfies $\sum_{k \geqslant 1}(1 - |z_k|) < \infty$. These *Blaschke products* are analytic on $\mathbb{D}$ and have the additional property that the radial limit $\lim_{r \to 1^-} B(re^{i\theta})$ exists and is of unit modulus for almost every $\theta \in [0, 2\pi)$. In other words, B is an inner function. Blaschke products have been studied intensely since they were first introduced and they appear in many contexts throughout complex analysis and operator theory.

This book is concerned with *finite Blaschke products*, in which the zero sequence $z_1, z_2, \ldots, z_n$ is finite and the product terminates. Although the skeptical reader might think this focus is too narrow, there are many fascinating connections with geometry, complex analysis, and operator theory that demand attention.

There are already some excellent texts that cover infinite Blaschke products and, more generally, inner functions [38, 61]. However, as the reader will see, there are many beautiful theorems involving finite Blaschke products that have no clear analogues in the infinite case. Finite Blaschke products are not often discussed in the standard texts on function spaces or complex variables since the focus there is often on inner functions as part of the broader theory of Hardy spaces. This book focuses on finite Blaschke products and the many results that pertain only to the finite case.

The book begins with an exposition of the *Schur class* $\mathscr{S}$, the set of analytic functions from $\mathbb{D}$ to $\mathbb{D}^-$, the closure of $\mathbb{D}$, and an introduction to hyperbolic geometry. We develop this material from scratch, assuming only that the reader has had a basic course in complex variables. We characterize the finite Blaschke products in several different ways. First, a rational function is a finite Blaschke product if and only if it is of the form

$$\frac{\alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n}{\overline{\alpha_n} + \overline{\alpha_{n-1}} z^{n-1} + \cdots + \overline{\alpha_0} z^n},$$

in which the numerator is a polynomial whose n roots lie in $\mathbb{D}$. Second, a finite Blaschke product maps $\mathbb{D}$ onto $\mathbb{D}$ (and the unit circle $\mathbb{T}$ onto itself) precisely n times and a theorem of Fatou confirms that these are the only functions that are continuous on $\mathbb{D}^-$ and analytic on $\mathbb{D}$ with this property. Third, each finite Blaschke product B satisfies

$$\lim_{|z| \to 1^-} |B(z)| = 1$$

and another result of Fatou shows that the finite Blaschke products are the only analytic functions on $\mathbb{D}$ that do this. Whether as rational functions whose defining polynomials enjoy certain symmetries, as n-to-1 analytic functions on $\mathbb{D}$, or as analytic functions with unimodular boundary values, the finite Blaschke products distinguish themselves as special elements of the Schur class.

The approximation of a given analytic function by well-understood functions from a fixed class is a standard technique in complex analysis. For example, there are the well-known approximation theorems of Runge, Mergelyan, and Weierstrass. We examine a few results of this type that involve finite Blaschke products. More specifically, a celebrated theorem of Carathéodory ensures that any function in the Schur class $\mathscr{S}$ can be approximated, uniformly on compact subsets of $\mathbb{D}$, by a sequence of finite Blaschke products. In fact, one can even take the approximating Blaschke products to have simple zeros. After Carathéodory's theorem, we discuss Fisher's theorem, which says that any function in $\mathscr{S}$ that extends continuously to $\mathbb{D}^-$ can be approximated uniformly on $\mathbb{D}^-$ by convex combinations of finite Blaschke products. As another example, a theorem of Helson and Sarason states that any continuous function from $\mathbb{T}$ to $\mathbb{T}$ can be uniformly approximated by a sequence of quotients of finite Blaschke products.

One might think there is not much to say about the zeros of a finite Blaschke product. After all, the location of the zeros is part of the definition! However, there are some beautiful gems here. The famed Gauss–Lucas theorem asserts that if P is a polynomial, then the zeros of P', the derivative of P, are contained in the convex hull of the zeros of P. There are theorems that say that the zeros of a finite Blaschke product B are contained in the convex hull of the solutions to the equation $B(z) = 1$ (or indeed the solutions to $B(z) = e^{i\theta}$ for any $\theta \in [0, 2\pi)$). Moreover, the hyperbolic analogue of the Gauss–Lucas theorem says that the zeros of B' (the critical points of B) are contained in the hyperbolic convex hull of the zeros of B. For Blaschke products of low degree, these results are even more explicit and can be stated in terms of classical geometry involving ellipses. There is also a result of Heins which says that one can create a finite Blaschke product with any desired set of critical points in $\mathbb{D}$. Finally, for analytic functions on $\mathbb{D}^-$, one can state, in terms of finite Blaschke products, a curious converse (the Challener–Rubel theorem) to Rouché's theorem.

Interpolation is another important topic in complex analysis. The most basic result in this direction is the Lagrange interpolation theorem, which guarantees that for distinct $z_1, z_2, \ldots, z_n$ and any $w_1, w_2, \ldots, w_n$ there is a polynomial P for which $P(z_j) = w_j$ for all j. The connection finite Blaschke products make with interpolation comes from Pick's theorem: given distinct $z_1, z_2, \ldots, z_n \in \mathbb{D}$ and any $w_1, w_2, \ldots, w_n \in \mathbb{D}$, then there is an $f \in \mathscr{S}$ for which $f(z_j) = w_j$ for all j if and only if the *Pick matrix*

$$\left[\frac{1 - \overline{w_j} w_i}{1 - \overline{z_j} z_i}\right]_{1 \leqslant i, j \leqslant n}$$

is positive semidefinite. Furthermore, when the interpolation is possible, it can be done with a finite Blaschke product. A more involved boundary interpolation result is the Cantor–Phelps theorem (for which we provide two distinct proofs, one abstract and another constructive), which says that given distinct $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ and any $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{T}$ there is a finite Blaschke product B with $B(\zeta_j) = \xi_j$ for all j.

So far we have discussed finite Blaschke product themselves and their connections to well-studied topics in complex analysis (zeros, critical points, residues, valence, approximation, and interpolation). However, as mentioned earlier, finite Blaschke products appear in many other places.

For example, Bohr's inequality asserts that if $f = \sum_{n \geqslant 0} a_n z^n \in \mathscr{S}$, then

$$\sum_{n \geqslant 0} |a_n| r^n \leqslant 1, \qquad r \in [0, \tfrac{1}{3}].$$

The number $\frac{1}{3}$ is optimal and is called the *Bohr radius* for the Schur class. Using finite Blaschke products, we explore a Bohr-type inequality for subclasses of Schur functions that vanish at certain points of $\mathbb{D}$ and for the Schur class functions whose first several derivatives vanish at zero. It turns out that the extremal functions for these extended Bohr problems are finite Blaschke products.

Next we cover two connections finite Blaschke products make with group theory. For a fixed finite Blaschke product B, consider the set G_B of continuous functions $u : \mathbb{T} \to \mathbb{T}$ for which $B \circ u = B$. One can see that G_B is a semigroup under function composition. A theorem of Chalendar and Cassier reveals that G_B is a cyclic group and that one can identify a generator by considering the previously mentioned n-to-1 mapping properties of B on $\mathbb{T}$. We also cover, via Cowen's unpublished exposition, an old theorem of Ritt that examines when we can write B as a composition $B = C \circ D$, in which C and D are finite Blaschke products. The answer is in terms of the monodromy group of B^{-1}. We also give an equivalent formulation of Ritt's theorem in terms of certain subgroups of G_B.

Finite Blaschke products also make connections to operator theory. For example, if T is a contraction on a Hilbert space and B is a finite Blaschke product with n zeros, then $B(T)$ is also a contraction. Moreover, a theorem of Gau and Wu says that $\|B(T)\| = 1$ if and only if $\|T^n\| = 1$. Another connection is with the numerical range of an operator. The spectral mapping theorem says that $\sigma(p(T)) = p(\sigma(T))$, in which $\sigma(T)$ is the spectrum of a bounded Hilbert space operator T and p is a polynomial. One may wonder whether or not a similar identity $W(p(T)) = p(W(T))$ holds for the *numerical range*

$$W(T) = \{\langle T\mathbf{x}, \mathbf{x}\rangle : \|\mathbf{x}\| = 1\}.$$

Although the desired identity is not true in general, there are some suitable substitutes. In fact, Halmos asked whether or not $W(T) \subseteq \mathbb{D}^-$ implies that $W(T^n) \subseteq \mathbb{D}^-$ for every $n \geqslant 1$. Progress was made when it was shown that if $W(T) \subseteq \mathbb{D}^-$ and B is a finite Blaschke product with $B(0) = 0$, then $W(B(T)) \subseteq \mathbb{D}^-$. A theorem of Berger and Stampfli extends this result from finite Blaschke products that vanish at the origin to the Schur functions that are continuous on $\mathbb{D}^-$ and vanish at the origin. However, without the condition $f(0) = 0$, there are contractions T with $W(T) \subseteq \mathbb{D}^-$ for which $W(f(T))$ intersects the complement of $\mathbb{D}^-$. A suitable replacement here is a theorem of Drury which says that though $W(f(T))$ may intersect the complement of $\mathbb{D}^-$, it is contained in a certain "teardrop" region, a slight "bulge" of $\mathbb{D}$. Moreover, the use of finite Blaschke products indicates the sharpness of Drury's theorem.

Still another connection to finite Blaschke products comes with models of linear transformations. In linear algebra, or more broadly in operator theory, one often wants to create a model for certain types of linear transformations. For example, there is the classical spectral theorem from linear algebra which says that any normal matrix is unitarily equivalent to a diagonal matrix. One can show that any contractive matrix T with $\operatorname{rank}(I - T^*T) = 1$ and whose eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are contained in $\mathbb{D}$ is unitarily equivalent to the compression of the shift operator $f \mapsto zf$ on the Hardy space H^2 to the model space

$$\operatorname{span}\left\{\frac{1}{1 - \overline{\lambda_j} z} : 1 \leqslant j \leqslant n\right\}.$$

Along with this result, one obtains a function-theoretic characterization of the invariant subspaces of these operators as well. In fact, this model space is the vector space of rational functions f with no poles in $\mathbb{D}^-$ for which

$$\int_0^{2\pi} f(e^{i\theta})\overline{B(e^{i\theta})}e^{-in\theta}\frac{d\theta}{2\pi} = 0, \qquad n \geqslant 0,$$

in which B is the finite Blaschke product whose zeros are the eigenvalues λ_j. The finite-dimensional approach undertaken in this book is intuitive and prepares interested readers for the more advanced text [59].

Finite Blaschke products can also be used to explore rational functions f that are analytic on $\mathbb{D}$ and for which $f(e^{i\theta})$ is an extended real number for all $\theta \in [0, 2\pi]$. These functions are sometimes called the real rational functions. Examples include

$$f(z) = i\frac{1+z}{1-z},$$

and, more generally,

$$f = i\frac{B_1 + B_2}{B_1 - B_2},$$

in which B_1 and B_2 are finite Blaschke products such that $B_1 - B_2$ has no zeros on $\mathbb{D}$. In fact, a theorem of Helson says these are all of the real rational functions. We will discuss various properties of real rational functions such as a characterization of those that are zero free on $\mathbb{D}$, the valence of these functions, as well as a factorization of a real rational function f as $f = FG$, where F and G are real rational functions, F has the same zeros of f, and G is zero free.

Finally, there is the connection Blaschke products make with the Darlington synthesis problem from electrical engineering. Here, in its simplest realization, one is given a rational function a with no poles in $\mathbb{D}^-$ and one needs to find rational functions b, c, d on $\mathbb{D}$ with no poles in $\mathbb{D}^-$ so that the matrix-valued analytic function

$$M(z) = \begin{bmatrix} a(z) & b(z) \\ c(z) & d(z) \end{bmatrix}$$

is such that $M(e^{i\theta})$ is a unitary matrix for every $\theta \in [0, 2\pi)$. The determinant of such a matrix M is a finite Blaschke product B and the model space associated with B determines the structure of and relations between the unknown functions b, c, d. Most curiously, we see that every rational matrix inner function $M(z)$ enjoys a peculiar quaternionic structure.

This book is mostly self-contained and should be accessible to a student with a background in basic real and complex analysis along with linear algebra. The proofs are detailed and dozens of illustrations are provided. We thank Zach Glassman for his assistance with `Tikz` and for producing many of our illustrations. At the end of each chapter, we include exercises so that the reader can gain greater technical fluency with the material. An appendix contains some background information about operator theory and function spaces that is relevant for a few results in the later chapters.

Claremont, CA, USA — Stephan Ramon Garcia
Laval, QC, Canada — Javad Mashreghi
Richmond, VA, USA — William T. Ross

Contents

Notation

$\delta_{j,k}$	Kronecker delta
$\mathbb{N} := \{1, 2, \ldots\}$	The set of natural numbers
$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers
$\mathbb{C}^n$	Complex n-space
A^*	Conjugate transpose of a matrix A
$\operatorname{Re} z$	Real part of a complex number z
$\operatorname{Im} z$	Imaginary part of a complex number z
$\equiv$	Identically equals
E^-	Closure of E
∂E	Boundary of E
$\lvert E\rvert$	Cardinality of a set E
$\operatorname{Res}(f, z_0)$	Residue of f at z_0
$\operatorname{diag}(z_1, \ldots, z_n)$	$n \times n$ diagonal matrix with $z_1, \ldots, z_n$ on the diagonal
$\mathbb{D} := \{z \in \mathbb{C} : \lvert z\rvert < 1\}$	Open unit disk (p. 1)
$\mathbb{D}^- := \{z \in \mathbb{C} : \lvert z\rvert \leqslant 1\}$	Closed unit disk (p. 1)
$\mathbb{T} := \{z \in \mathbb{C} : \lvert z\rvert = 1\}$	Unit circle (p. 1)
$\mathscr{S}$	Schur class (p. 1)
$\operatorname{Aut}(\mathbb{D})$	The automorphism group of $\mathbb{D}$ (p. 2)
id	Identity function (p. 2)
$\rho_\gamma(z) = \gamma z$	A rotation by $\arg \gamma$ (p. 2)
$\tau_w(z) = (w - z)/(1 - \overline{w}z)$	A special disk automorphism (p. 2)
$\widehat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$	Riemann sphere (p. 7)
$S_C(\alpha)$	Stolz domain anchored at α with constant C (p. 12)
$\angle \lim_{z\to\alpha} f(z)$	Nontangential limit (p. 12)
$\mathbb{D}_- := \mathbb{D} \cap \{z : \operatorname{Im} z < 0\}$	Lower half of unit disk (p. 16)
$\mathbb{D}_+ := \mathbb{D} \cap \{z : \operatorname{Im} z > 0\}$	Upper half of unit disk (p. 16)
$\mathbb{C}_+ := \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$	Upper half plane (p. 17)
$\mathbb{C}_- := \{z \in \mathbb{C} : \operatorname{Im} z < 0\}$	Lower half plane (p. 17)

d	Euclidean metric on $\mathbb{C}$ (p. 21)
ϱ	Pseudohyperbolic metric (p. 21)
$\Delta(z, \rho)$	Pseudohyperbolic disk of radius ρ centered at z (p. 22)
$D(c, r)$	Euclidean disk of radius r centered at c (p. 22)
$\wp$	Poincaré metric (p. 27)
$\ell(\Gamma)$	Hyperbolic length of a curve Γ (p. 27)
d_μ	Metric induced by μ (p. 31)
$\triangle$	Laplace operator (p. 32)
κ_μ	Curvature of the metric μ (p. 32)
$\mathbb{D}_e = \{z \in \mathbb{C} : \|z\| > 1\} \cup \{\infty\}$	The complement of $\mathbb{D}^-$ in $\widehat{\mathbb{C}}$ (p. 39)
deg	Degree of a rational functional (p. 43)
$P^\#(z) := z^n \overline{P(1/\overline{z})}$	In which P is a polynomial of degree n (p. 44)
$P^{\#_n}$	$P^\#$ when n needs to be specified (p. 44)
$\mathscr{A}(\mathbb{D})$	The disk algebra (p. 49)
H^∞	The bounded analytic functions on $\mathbb{D}$ (p. 59)
$\\| \cdot \\|_\infty$	Norm on H^∞ (p. 59)
conv	Convex hull (p. 61)
$Z_f(\Gamma)$	Number of zeros of f inside a curve Γ (p. 69)
$P_f(\Gamma)$	Number of poles of f inside a curve Γ (p. 69)
$\beta(a, z)$	(p. 109)
$\mathscr{B}_d$	(p. 110)
Σ_d	(p. 112)
Φ	(p. 114)
R_f	The distance-ratio function for f (p. 116)
$\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$	Extended real line (p. 131)
M_n	the set of $n \times n$ complex matrices (p. 135)
$\mathfrak{B}_0$	Bohr radius (p. 156)
$m(f, r)$	(p. 156)
$M(\mathscr{F}, r)$	(p. 157)
$\mathfrak{B}_k$	The kth Bohr radius (p. 163)
$\mathfrak{B}_k(\lambda)$	A generalized Bohr radius (p. 165)
$\mathscr{S}_\lambda$	(p. 165)
G_B	(p. 183)
$\mathscr{S}_B$	Critical values of a Blaschke product (p. 183)
$\widetilde{B}$	Normalized form of a finite Blaschke product B (p. 186)
$\mathscr{G}_B$	Monodromy group for a Blaschke product (p. 191)
S_n	Group of permutations of $\{1, 2, \ldots, n\}$ (p. 192)
$\mathrm{Hol}(\mathbb{D}^-)$	The set of functions analytic on a neighborhood of $\mathbb{D}^-$ (p. 215)
$\sigma(T)$	Spectrum of an operator T (p. 219)
$W(T)$	Numerical range of an operator T (p. 219)
$w(T)$	Numerical radius of an operator T (p. 220)
$\mathrm{td}(\alpha)$	Drury's teardrop region (p. 235)

$\mathfrak{R}$	The set rational of functions with real boundary values (p. 245)
$\mathfrak{R}^+$	The set of $\mathfrak{R}$ functions which are analytic on $\mathbb{D}$ (p. 245)
$\mathscr{K}_B$	Model space associated with a Blaschke product B (p. 261)
$b_w = (z - w)/(1 - \overline{w}z)$	a special disk automorphism (p. 267)
$k_\lambda(z)$	Reproducing kernel for $\mathscr{K}_B$ (p. 268)
$c_\lambda(z)$	Cauchy kernel (p. 268)
$\widetilde{c}_\lambda(z)$	Normalized Cauchy kernel (p. 268)
S_B	Compressed shift (p. 276)
i_F	Finite Blaschke product associated with F (p. 294)
$\\|T\\|$	Operator norm of an operator T (p. 313)
$\oplus$	Orthogonal direct sum (p. 314)
T_ϕ	Toeplitz operator with symbol ϕ (p. 316)

Chapter 1
Geometry of the Schur Class

This chapter will cover some basic facts about the Schur class. In what follows,

$$\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}, \quad \mathbb{D}^- = \{z \in \mathbb{C} : |z| \leqslant 1\}, \quad \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}.$$

Definition 1.0.1 The *Schur class* $\mathscr{S}$ is

$$\mathscr{S} := \{f : \mathbb{D} \to \mathbb{D}^- : \ f \text{ is analytic}\}. \tag{1.0.2}$$

The Maximum Modulus Principle ensures that $f(z) \in \mathbb{T}$ for some $z \in \mathbb{D}$ if and only if f is a constant function of unit modulus. Thus, $\mathscr{S}$ consists of the nonconstant analytic functions $f : \mathbb{D} \to \mathbb{D}$ along with the constant functions with values in $\mathbb{D}^-$.

1.1 The Schwarz Lemma

The Schwarz Lemma is one of the cornerstones of complex analysis. Despite its deceptive simplicity, it has many profound consequences [31]. Schwarz proved this lemma for injective functions. Carathéodory proved the general version.

Lemma 1.1.1 (Schwarz [125]) *If $f \in \mathscr{S}$ and $f(0) = 0$, then*

(a) $|f(z)| \leqslant |z|$ *for all* $z \in \mathbb{D}$, *and*
(b) $|f'(0)| \leqslant 1$.

Moreover, if $|f(w)| = |w|$ for some $w \in \mathbb{D}\backslash\{0\}$ or if $|f'(0)| = 1$, then there is a $\zeta \in \mathbb{T}$ so that $f(z) = \zeta z$ for all $z \in \mathbb{D}$.

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_1

Proof (Carathéodory [15]) Define $g : \mathbb{D} \to \mathbb{C}$ by

$$g(z) = \begin{cases} \dfrac{f(z)}{z} & \text{if } z \neq 0, \\ f'(0) & \text{if } z = 0, \end{cases}$$

and observe that g is analytic on $\mathbb{D}\backslash\{0\}$. The singularity at 0 is removable since

$$\lim_{z \to 0} g(z) = f'(0)$$

and hence g is analytic on all of $\mathbb{D}$. For $r \in [0, 1)$, an application of the Maximum Modulus Principle to the disk $|z| \leqslant r$ yields a $\zeta \in \mathbb{T}$ so that

$$|g(rz)| \leqslant |g(r\zeta)| = \frac{|f(r\zeta)|}{|r\zeta|} \leqslant \frac{1}{r}, \quad z \in \mathbb{D}.$$

Now let $r \to 1^-$ to obtain statements (a) and (b).

Suppose that $|f(w)| = |w|$ for some $w \in \mathbb{D}\backslash\{0\}$ or that $|f'(0)| = 1$. Then $|g(w)| = 1$ for some $w \in \mathbb{D}$. Since $|g| \leqslant 1$ on $\mathbb{D}$, the Maximum Modulus Principle provides a $\zeta \in \mathbb{T}$ such that $g(z) = \zeta$ for all $z \in \mathbb{D}$. Thus, $f(z) = \zeta z$ for all $z \in \mathbb{D}$. □

1.2 Automorphisms of the Disk

Definition 1.2.1 A bijective analytic function $f : \mathbb{D} \to \mathbb{D}$ is an *automorphism* of $\mathbb{D}$.

Since most of our work concerns the unit disk $\mathbb{D}$, we often say "f is an automorphism" without explicit reference to $\mathbb{D}$. The set of all automorphisms of $\mathbb{D}$, denoted by $\operatorname{Aut}(\mathbb{D})$, is a subset of the Schur class $\mathscr{S}$.

If f is an automorphism, then the inverse bijection $f^{-1} : \mathbb{D} \to \mathbb{D}$ is analytic and hence f^{-1} is also an automorphism. The *identity function* $\mathrm{id} : \mathbb{D} \to \mathbb{D}$ defined by

$$\mathrm{id}(z) = z$$

is an automorphism satisfying $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}$ for every $f \in \operatorname{Aut}(\mathbb{D})$. Since the composition of two automorphisms is also an automorphism, and since function composition is an associative operation, $\operatorname{Aut}(\mathbb{D})$ is a group under function composition.

We now focus on two special automorphisms. For $w \in \mathbb{D}$ and $\gamma \in \mathbb{T}$, define $\rho_\gamma : \mathbb{D} \to \mathbb{C}$ and $\tau_w : \mathbb{D} \to \mathbb{C}$ by

$$\rho_\gamma(z) = \gamma z \qquad \text{and} \qquad \tau_w(z) = \frac{w - z}{1 - \overline{w}z}. \tag{1.2.2}$$

Since $|\gamma| = 1$, we see that ρ_γ induces a rotation of $\mathbb{D}$ about the origin through an angle of $\arg\gamma$. Consequently, $\rho_\gamma \in \mathrm{Aut}(\mathbb{D})$. Moreover,

$$\rho_{\gamma_1} \circ \rho_{\gamma_2} = \rho_{\gamma_1\gamma_2} \qquad \text{and} \qquad \rho_\gamma \circ \rho_{\overline{\gamma}} = \mathrm{id}. \tag{1.2.3}$$

The function τ_w is also an automorphism of $\mathbb{D}$, although to establish this requires a little more work. First, a computation confirms that

$$\tau_w \circ \tau_w = \mathrm{id}, \tag{1.2.4}$$

so τ_w is injective on $\mathbb{D}$ and the range of τ_w contains $\mathbb{D}$. To show that the range of τ_w is precisely $\mathbb{D}$, observe that for each $\zeta \in \mathbb{T}$ and $w \in \mathbb{D}$,

$$|\tau_w(\zeta)| = \left|\frac{w - \zeta}{1 - \overline{w}\zeta}\right| = \frac{|w - \zeta|}{|\overline{w} - \overline{\zeta}|} = 1$$

since $\zeta\overline{\zeta} = |\zeta|^2 = 1$. Since the Maximum Modulus Principle implies that

$$|\tau_w(z)| < 1, \qquad z \in \mathbb{D},$$

it follows that $\tau_w \in \mathrm{Aut}(\mathbb{D})$. Therefore, by the discussion above,

$$\{\rho_\gamma \circ \tau_w : \gamma \in \mathbb{T}, w \in \mathbb{D}\} \subseteq \mathrm{Aut}(\mathbb{D}).$$

The following theorem establishes that the preceding containment is an equality.

Theorem 1.2.5 *If* $f \in \mathrm{Aut}(\mathbb{D})$, *then there are unique* $w \in \mathbb{D}$ *and* $\gamma \in \mathbb{T}$ *such that* $f = \rho_\gamma \circ \tau_w$. *In other words,*

$$\mathrm{Aut}(\mathbb{D}) = \{\rho_\gamma \circ \tau_w : \gamma \in \mathbb{T},\ w \in \mathbb{D}\}.$$

Proof If $f \in \mathrm{Aut}(\mathbb{D})$, then there is a unique $w \in \mathbb{D}$ so that $f(w) = 0$. Then $g = f \circ \tau_w \in \mathrm{Aut}(\mathbb{D})$ and $g(0) = 0$. Hence the Schwarz Lemma (Lemma 1.1.1) ensures that

$$|g(z)| \leqslant |z|, \qquad z \in \mathbb{D}.$$

Since $g^{-1} \in \mathrm{Aut}(\mathbb{D})$ and $g^{-1}(0) = 0$, the same argument yields

$$|g^{-1}(z)| \leqslant |z|, \qquad z \in \mathbb{D}.$$

Since $g(z) \in \mathbb{D}$, we may substitute $g(z)$ in place of z in the previous inequality and obtain $|z| \leqslant |g(z)|$ for all $z \in \mathbb{D}$. Consequently,

$$|g(z)| = |z|, \qquad z \in \mathbb{D},$$

and hence another application of the Schwarz Lemma yields a unique unimodular constant γ such that $g(z) = \gamma z$. Thus, $f(\tau_w(z)) = \gamma z$ for all $z \in \mathbb{D}$. Now substitute z in place of $\tau_w(z)$ in the preceding identity and use (1.2.4) to obtain $f = \gamma\tau_w = \rho_\gamma \circ \tau_w$.

We now verify the uniqueness of the parameters γ and w in the representation $\rho_\gamma \circ \tau_w$ of a typical element of Aut($\mathbb{D}$). Suppose that

$$\rho_\gamma \circ \tau_w = \rho_{\gamma'} \circ \tau_{w'}$$

for some $\gamma, \gamma' \in \mathbb{T}$ and $w, w' \in \mathbb{D}$. Then (1.2.3) and (1.2.4) yield

$$\rho_{\gamma\overline{\gamma'}} = \tau_{w'} \circ \tau_w.$$

Evaluate the preceding identity at $z = 0$ to obtain $\tau_{w'}(w) = 0$ and so $w = w'$. Hence $\rho_{\gamma\overline{\gamma'}} = \mathrm{id}$ and thus $\gamma = \gamma'$. □

Since $\tau_0 = -\mathrm{id}$ and $\rho_1 = \mathrm{id}$, the unique representations of τ_w and ρ_γ afforded by Theorem 1.2.5 are

$$\tau_w = \rho_1 \circ \tau_w$$

and

$$\rho_\gamma = \rho_{-\gamma} \circ \tau_0. \tag{1.2.6}$$

It is also worth noting that if $f \in \mathrm{Aut}(\mathbb{D})$ and $f(0) = 0$, then $f = \rho_\gamma$ for some $\gamma \in \mathbb{T}$; that is, the only automorphisms of $\mathbb{D}$ that fix the origin are the rotations.

1.3 Algebraic Structure of Aut($\mathbb{D}$)

If $f = \rho_{\gamma_1} \circ \tau_{w_1}$ and $g = \rho_{\gamma_2} \circ \tau_{w_2}$ are automorphisms of $\mathbb{D}$, then Theorem 1.2.5 implies that $f \circ g = \rho_\gamma \circ \tau_w$ for some unique $\gamma \in \mathbb{T}$ and $w \in \mathbb{D}$. Since we often require concrete formulas that are applicable to problems in function theory, our primary goal in this section is to obtain expressions for γ and w in terms of the parameters γ_1, γ_2, w_1, and w_2. At the end of this section, however, we will briefly describe a more group-theoretic approach to Aut($\mathbb{D}$).

Lemma 1.3.1 *If $f = \rho_\gamma \circ \tau_w$, then $w = f^{-1}(0)$ and*

$$\gamma = \begin{cases} f(0)/f^{-1}(0) & \text{if } f(0) \neq 0, \\ -f'(0) & \text{if } f(0) = 0. \end{cases}$$

Proof Since

$$f(w) = \rho_\gamma(\tau_w(w)) = \rho_\gamma(0) = 0$$

and f is invertible, we conclude that $w = f^{-1}(0)$. Moreover,

$$f(0) = \rho_\gamma(\tau_w(0)) = \rho_\gamma(w) = \gamma w = \gamma f^{-1}(0),$$

which yields the desired formula when $f(0) \neq 0$. When $f(0) = 0$, we get

$$w = f^{-1}(0) = 0$$

and hence

$$f(z) = \rho_\gamma(\tau_0(z)) = \rho_\gamma(-z) = -\gamma z.$$

Thus, $\gamma = -f'(0)$ as claimed. □

The discussion below requires the following derivative formula:

$$\tau_w'(z) = -\frac{1-|w|^2}{(1-\overline{w}z)^2}.$$

Let $z = 0$ and $z = w$, respectively, in the preceding and obtain

$$\tau_w'(0) = -(1-|w|^2) \tag{1.3.2}$$

and

$$\tau_w'(w) = -\frac{1}{1-|w|^2}. \tag{1.3.3}$$

The following theorem provides an explicit realization of the group operation on Aut($\mathbb{D}$). It also yields several formulas that are needed later on.

Theorem 1.3.4 *If $\gamma_1, \gamma_2 \in \mathbb{T}$ and $w_1, w_2 \in \mathbb{D}$, then*

$$\big(\rho_{\gamma_1} \circ \tau_{w_1}\big) \circ \big(\rho_{\gamma_2} \circ \tau_{w_2}\big) = \rho_\gamma \circ \tau_w,$$

where

$$w = \tau_{w_2}(\overline{\gamma_2} w_1)$$

and

$$\gamma = \begin{cases} \gamma_1 \tau_{w_1\overline{w_2}}(\gamma_2) & \text{if } w_2 \neq \overline{\gamma_2} w_1, \\ -\gamma_1\gamma_2 & \text{if } w_2 = \overline{\gamma_2} w_1. \end{cases}$$

In particular, if $w_2 = \overline{\gamma_2} w_1$, then

$$\big(\rho_{\gamma_1} \circ \tau_{w_1}\big) \circ \big(\rho_{\gamma_2} \circ \tau_{w_2}\big) = \rho_{\gamma_1\gamma_2}.$$

Proof Let $f = (\rho_{\gamma_1} \circ \tau_{w_1}) \circ (\rho_{\gamma_2} \circ \tau_{w_2})$. Lemma 1.3.1 says that w is the unique solution to the equation

$$f(w) = \big[(\rho_{\gamma_1} \circ \tau_{w_1}) \circ (\rho_{\gamma_2} \circ \tau_{w_2})\big](w) = 0.$$

Since $\rho_{\gamma_1}(0) = 0$, we see that

$$\left[\tau_{w_1} \circ (\rho_{\gamma_2} \circ \tau_{w_2})\right](w) = 0$$

and hence

$$\left(\rho_{\gamma_2} \circ \tau_{w_2}\right)(w) = \tau_{w_1}(0) = w_1$$

by (1.2.4). An application of (1.2.3) yields

$$\tau_{w_2}(w) = \rho_{\overline{\gamma_2}}(w_1) = \overline{\gamma_2} w_1, \tag{1.3.5}$$

after which another appeal to (1.2.4) provides the desired formula for w. Now observe that the preceding formula yields

$$w = 0 \quad \Longleftrightarrow \quad w_2 = \overline{\gamma_2} w_1.$$

Since $w = f^{-1}(0)$, the second formula in Lemma 1.3.1 asserts that $\gamma = f(0)/w$ when $w \neq 0$. The computation

$$\begin{aligned} f(0) &= \left[(\rho_{\gamma_1} \circ \tau_{w_1}\right) \circ \left(\rho_{\gamma_2} \circ \tau_{w_2})\right](0) \\ &= \gamma_1 \tau_{w_1}\left(\gamma_2 \tau_{w_2}(0)\right) \\ &= \gamma_1 \tau_{w_1}(\gamma_2 w_2) \end{aligned}$$

and (1.3.5) reveal that

$$\gamma = \frac{f(0)}{w} = \frac{\gamma_1 \tau_{w_1}\left(\gamma_2 w_2\right)}{\tau_{w_2}(\overline{\gamma_2} w_1)} = \gamma_1 \tau_{w_1 \overline{w_2}}(\gamma_2).$$

The final equality in the statement of the theorem is verified by direct computation.

If $w = 0$, then we need to evaluate $f'(0)$. By the chain rule and (1.3.2),

$$\begin{aligned} f'(0) &= \gamma_1 \tau'_{w_1}[(\rho_{\gamma_2} \circ \tau_{w_2})(0)] \times \gamma_2 \tau'_{w_2}(0) \\ &= -\gamma_1 \tau'_{w_1}(\gamma_2 w_2) \times \gamma_2(1 - |w_2|^2) \\ &= -\gamma_1 \gamma_2. \end{aligned}$$

□

Corollary 1.3.6 *If $w_1, w_2 \in \mathbb{D}$ and $w_1 \neq w_2$, then*

$$\tau_{w_1} \circ \tau_{w_2} = \rho_\gamma \circ \tau_w,$$

where

$$w = \tau_{w_2}(w_1) = \frac{w_2 - w_1}{1 - \overline{w_2} w_1} \quad \textit{and} \quad \gamma = \tau_{w_1 \overline{w_2}}(1) = -\frac{1 - w_1 \overline{w_2}}{1 - \overline{w_1} w_2}.$$

For the following result, let $\gamma_1 = 1$ and $w_2 = 0$, then replace γ_2 by $-\gamma$ and w_1 by w in Theorem 1.3.4. However, we admit that direct verification might be easier; see Exercise 1.1.

Corollary 1.3.7 *If $w \in \mathbb{D}$ and $\gamma \in \mathbb{T}$, then*

$$\tau_w \circ \rho_\gamma = \rho_\gamma \circ \tau_{\overline{\gamma} w}.$$

Although Theorem 1.3.4 provides an explicit description, in terms of the factorization afforded by Theorem 1.2.5, of the group operation on Aut($\mathbb{D}$), an algebraist might find our approach unsatisfactory. Let us briefly discuss a more abstract approach to Aut($\mathbb{D}$).

A *Möbius transformation* (also called a *linear fractional transformation*) is a rational function of the form

$$f(z) = \frac{az + b}{cz + d}, \tag{1.3.8}$$

in which $ad - bc \neq 0$. Each Möbius transformation is a bijective map from the *extended complex plane* $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ (or *Riemann sphere*) to itself. The set of all Möbius transformations is a group under composition; the identity is the function $\mathrm{id}(z) = z$ and the inverse of f is

$$f^{-1}(z) = \frac{dz - b}{-cz + a}.$$

If we multiply the numerator and denominator of (1.3.8) by a suitable constant, we may assume that $ad - bc = 1$.

The group of Möbius transformations is isomorphic to $\mathrm{PSL}_2(\mathbb{C})$, the *projective special linear group of order* 2 *over* $\mathbb{C}$. To be more specific, $\mathrm{PSL}_2(\mathbb{C})$ is the quotient of $\mathrm{SL}_2(\mathbb{C})$, the group of 2×2 complex matrices with determinant 1, by the subgroup $\{I, -I\}$. Here I denotes the 2×2 identity matrix. The isomorphism between the group of Möbius transformations and $\mathrm{PSL}_2(\mathbb{C})$ is given by sending the function in (1.3.8), in which $ad - bc = 1$, to the coset of

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

in $\mathrm{SL}_2(\mathbb{C})/\{I, -I\}$.

Theorem 1.2.5 asserts that $\mathrm{Aut}(\mathbb{D}) = \{\rho_\gamma \circ \tau_w : \gamma \in \mathbb{T}, w \in \mathbb{D}\}$, in which

$$\rho_\gamma(z) = \frac{\gamma z + 0}{0z + 1} \quad \text{and} \quad \tau_w(z) = \frac{-1z + w}{-\overline{w}z + 1}.$$

The cosets in $\mathrm{SL}_2(\mathbb{C})/\{I, -I\}$ that correspond to ρ_γ and τ_w are the cosets of

$$\begin{bmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \alpha & \beta \\ \overline{\beta} & \overline{\alpha} \end{bmatrix},$$

where

$$\gamma = e^{i\theta}, \qquad \alpha = \frac{i}{\sqrt{1-|w|^2}}, \quad \text{and} \quad \beta = \frac{-iw}{\sqrt{1-|w|^2}}.$$

Consequently, $\mathrm{Aut}(\mathbb{D})$ can be identified with $\mathrm{PSU}_{1,1}(\mathbb{C})$, the quotient of

$$\mathrm{SU}_{1,1}(\mathbb{C}) = \left\{ \begin{bmatrix} a & b \\ \overline{b} & \overline{a} \end{bmatrix} : |a|^2 - |b|^2 = 1 \right\}$$

by the subgroup $\{I, -I\}$. It is worth remarking that $\mathrm{SU}_{1,1}(\mathbb{C})$ is the set of 2×2 complex matrices U for which $\det U = 1$ and $U^* \Gamma U = \Gamma$, in which U^* denotes the conjugate transpose of U and

$$\Gamma = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

This suggests a connection between $\mathrm{Aut}(\mathbb{D})$ and hyperbolic geometry that will be explored further in Chap. 2.

From a topological perspective, $\mathrm{Aut}(\mathbb{D})$ is homeomorphic to $\mathbb{T} \times \mathbb{D}$ via the map

$$(\gamma, w) \mapsto \rho_\gamma \circ \tau_w, \qquad \gamma \in \mathbb{T},\ w \in \mathbb{D}.$$

Thus, $\mathrm{Aut}(\mathbb{D})$ can be visualized as an open solid torus, endowed with the group structure described in Theorem 1.3.4.

1.4 The Schwarz–Pick Theorem

The hypothesis of the Schwarz Lemma (Lemma 1.1.1) involves a function that vanishes at the origin. A generalization can be obtained that removes this hypothesis. The crucial idea is to employ suitable automorphisms to reduce the general case to the classical Schwarz Lemma.

Theorem 1.4.1 (Schwarz–Pick) *For each $f \in \mathscr{S}$,*

$$\left| \frac{f(z) - f(w)}{1 - \overline{f(w)} f(z)} \right| \leqslant \left| \frac{z - w}{1 - \overline{w} z} \right|, \qquad w, z \in \mathbb{D}, \tag{1.4.2}$$

and

$$\frac{|f'(z)|}{1 - |f(z)|^2} \leqslant \frac{1}{1 - |z|^2}, \qquad z \in \mathbb{D}. \tag{1.4.3}$$

Moreover, the following are equivalent.

(a) Equality holds in (1.4.2) *at two distinct* $z, w \in \mathbb{D}$.
(b) Equality holds in (1.4.2) *at all* $z, w \in \mathbb{D}$ *with* $z \neq w$.
(c) Equality holds in (1.4.3) *at some* $z \in \mathbb{D}$.
(d) Equality holds in (1.4.3) *at all* $z \in \mathbb{D}$.
(e) $f \in \operatorname{Aut}(\mathbb{D})$.

Proof Fix $w \in \mathbb{D}$. If $|f(w)| = 1$, the Maximum Modulus Principle implies that f is constant which means that (1.4.2) and (1.4.3) hold automatically. On the other hand, if $f(w) \in \mathbb{D}$, the Maximum Modulus Principle implies that $f(\mathbb{D}) \subseteq \mathbb{D}$. Let

$$g = \tau_{f(w)} \circ f \circ \tau_w \tag{1.4.4}$$

and observe that $g : \mathbb{D} \to \mathbb{D}$ is analytic and $g(0) = 0$. Since

$$g(\tau_w(z)) = \frac{f(w) - f(z)}{1 - \overline{f(w)} f(z)} \qquad \text{and} \qquad g'(0) = \frac{1 - |z|^2}{1 - |f(z)|^2} f'(z),$$

we see that (1.4.2) is equivalent to

$$|g(\tau_w(z))| \leqslant |\tau_w(z)|, \qquad w, z \in \mathbb{D} \tag{1.4.5}$$

and (1.4.3) is equivalent to

$$|g'(0)| \leqslant 1. \tag{1.4.6}$$

However, (1.4.5) and (1.4.6) hold by the Schwarz Lemma.

If any of (a)–(d) hold, then an application of the Schwarz Lemma to g confirms that $g = \rho_\gamma$ for some $\gamma \in \mathbb{T}$. Thus, (1.4.4) ensures that $f \in \operatorname{Aut}(\mathbb{D})$. Conversely, if $f \in \operatorname{Aut}(\mathbb{D})$, then (1.4.4) implies that $g \in \operatorname{Aut}(\mathbb{D})$ with $g(0) = 0$ and thus $g = \rho_\gamma$ for some $\gamma \in \mathbb{T}$. For this automorphism g, equality holds in (1.4.5) and (1.4.6) and consequently equality holds in (1.4.2) and (1.4.3). In other words, (e) implies any of (a)–(d). □

As a special case of Theorem 1.4.1, let $f = \tau_{z_0}$ to obtain

$$\left| \frac{\tau_{z_0}(z) - \tau_{z_0}(w)}{1 - \overline{\tau_{z_0}(w)} \tau_{z_0}(z)} \right| = \left| \frac{z - w}{1 - \overline{w} z} \right|, \qquad z, w \in \mathbb{D}, \tag{1.4.7}$$

and

$$|\tau_{z_0}'(z)| = \frac{1 - |\tau_{z_0}(z)|^2}{1 - |z|^2}, \quad z \in \mathbb{D}. \tag{1.4.8}$$

These two identities will be useful later.

1.5 An Extremal Problem

Theorem 1.4.1 can be applied to solve certain extremal problems for $\mathscr{S}$. We briefly discuss one of them. Fix $\alpha, \beta \in \mathbb{D}$ and let

$$\mathscr{A}_{\alpha,\beta} = \{f \in \mathscr{S} : f(\alpha) = \beta\}.$$

Observe that $f = \tau_\beta \circ \tau_\alpha \in \mathscr{A}_{\alpha,\beta}$ and hence $\mathscr{A}_{\alpha,\beta} \neq \varnothing$. Our goal is to compute

$$M = \sup_{f \in \mathscr{A}_{\alpha,\beta}} |f'(\alpha)|,$$

along with the functions $f \in \mathscr{A}_{\alpha,\beta}$ for which the supremum above is attained.

Theorem 1.4.1 implies that

$$|f'(\alpha)| \leqslant \frac{1 - |f(\alpha)|^2}{1 - |\alpha|^2} = \frac{1 - |\beta|^2}{1 - |\alpha|^2}, \qquad f \in \mathscr{A}_{\alpha,\beta}.$$

A computation using (1.3.2) and (1.3.3) confirms that equality is attained when $f = \tau_\beta \circ \tau_\alpha$. Thus,

$$M = \frac{1 - |\beta|^2}{1 - |\alpha|^2}.$$

Moreover, Theorem 1.4.1 asserts that the $f \in \mathscr{A}_{\alpha,\beta}$ for which

$$|f'(\alpha)| = \frac{1 - |\beta|^2}{1 - |\alpha|^2}$$

are precisely the $f \in \operatorname{Aut}(\mathbb{D})$ that satisfy $f(\alpha) = \beta$. Let f be such an automorphism and let $g = \tau_\beta \circ f \circ \tau_\alpha$; observe that $g \in \operatorname{Aut}(\mathbb{D})$. Then

$$g(0) = \tau_\beta(f(\tau_\alpha(0))) = \tau_\beta(f(\alpha)) = \tau_\beta(\beta) = 0$$

and hence $g(z) = \gamma z$ for some $\gamma \in \mathbb{T}$; that is, $g = \rho_\gamma$. Hence the solutions to the extremal problem are given by

$$f = \tau_\beta \circ \rho_\gamma \circ \tau_\alpha,$$

in which $\gamma \in \mathbb{T}$ is a free parameter.

1.6 Julia's Lemma

The Schwarz–Pick theorem (Theorem 1.4.1) involves two points $z, w \in \mathbb{D}$. What happens if one of the points approaches $\mathbb{T}$? This situation was studied by Julia and it may be interpreted as a boundary Schwarz–Pick theorem [83, p. 87]. Julia's lemma plays an essential role in studying the behavior of the derivative of infinite Blaschke products. The proof of Julia's lemma requires the important identity

$$1 - \left| \frac{\alpha - \beta}{1 - \overline{\beta}\alpha} \right|^2 = \frac{(1 - |\alpha|^2)(1 - |\beta|^2)}{|1 - \overline{\beta}\alpha|^2}, \qquad \alpha, \beta \in \mathbb{D}, \tag{1.6.1}$$

which follows from (1.4.8).

Lemma 1.6.2 (Julia [83]) *Let $f \in \mathscr{S}$. If there is a sequence z_n in $\mathbb{D}$ such that*

$$\lim_{n\to\infty} z_n = 1, \qquad \lim_{n\to\infty} f(z_n) = 1,$$

and

$$\lim_{n\to\infty} \frac{1 - |f(z_n)|}{1 - |z_n|} = A < \infty, \tag{1.6.3}$$

then

$$\frac{|1 - f(z)|^2}{1 - |f(z)|^2} \leqslant A \frac{|1 - z|^2}{1 - |z|^2}, \qquad z \in \mathbb{D}. \tag{1.6.4}$$

Proof The Schwarz–Pick theorem (Theorem 1.4.1) implies that

$$\left| \frac{f(z) - f(z_n)}{1 - \overline{f(z_n)} f(z)} \right| \leqslant \left| \frac{z - z_n}{1 - \overline{z}_n z} \right|, \qquad z \in \mathbb{D},$$

and hence

$$1 - \left| \frac{z - z_n}{1 - \overline{z}_n z} \right|^2 \leqslant 1 - \left| \frac{f(z) - f(z_n)}{1 - \overline{f(z_n)} f(z)} \right|^2.$$

The identity (1.6.1), applied to both sides of the above, yields

$$\frac{(1 - |z|^2)(1 - |z_n|^2)}{|1 - \overline{z_n} z|^2} \leqslant \frac{(1 - |f(z)|^2)(1 - |f(z_n)|^2)}{|1 - \overline{f(z_n)} f(z)|^2}.$$

Rewrite the preceding inequality as

$$\frac{|1 - \overline{f(z_n)} f(z)|^2}{1 - |f(z)|^2} \leqslant \frac{1 - |f(z_n)|^2}{1 - |z_n|^2} \cdot \frac{|1 - \overline{z_n} z|^2}{1 - |z|^2}.$$

Now let $n \to \infty$ and apply (1.6.3) to complete the proof. □

In the lemma above, we assumed that $z_n \to 1$ and $f(z_n) \to 1$. However, the important issue is that the sequences z_n and $f(z_n)$ converge toward points of the unit circle $\mathbb{T}$. For the sake of completeness, here is the general version of this result.

Corollary 1.6.5 *Let $f \in \mathscr{S}$ and $\alpha, \beta \in \mathbb{T}$. If there is a sequence z_n in $\mathbb{D}$ such that*

$$\lim_{n\to\infty} z_n = \alpha, \qquad \lim_{n\to\infty} f(z_n) = \beta,$$

and

$$\lim_{n\to\infty} \frac{1-|f(z_n)|}{1-|z_n|} = A < \infty,$$

then

$$\frac{|\beta - f(z)|^2}{1-|f(z)|^2} \leqslant A\frac{|\alpha - z|^2}{1-|z|^2}, \qquad z \in \mathbb{D}.$$

Proof Apply Lemma 1.6.2 to the function $g(z) = \overline{\beta} f(\overline{\alpha} z)$. □

We can also discuss the boundary limits of functions in $\mathscr{S}$ that satisfy the hypotheses of Julia's Lemma. Let $\zeta \in \mathbb{T}$ and $C > 1$. The region

$$S_C(\zeta) = \{z \in \mathbb{D} : |z - \zeta| \leqslant C(1-|z|)\}$$

is the *Stolz domain* anchored at α with constant C; see Fig. 1.1.

We say that $f \in \mathscr{S}$ has the *nontangential limit* L at $\zeta \in \mathbb{T}$ if, for each fixed $C > 1$,

$$\lim_{\substack{z\to\zeta \\ z\in S_C(\zeta)}} f(z) = L. \tag{1.6.6}$$

If so, we define $f(\zeta) = L$ and write

$$\angle \lim_{z\to\zeta} f(z) = f(\zeta).$$

The quantity $f(\zeta)$ is referred to as *the boundary value* of f at ζ. The restriction that z belongs to a Stolz domain $S_C(\zeta)$ in (1.6.6) ensures that z does not approach ζ along a path that is tangent to $\mathbb{T}$ at ζ. Each Schur function has non-tangential boundary values almost everywhere with respect to Lebesgue measure on $\mathbb{T}$; see Theorem A.3.1.

Corollary 1.6.7 *Let $f \in \mathscr{S}$ and let $\alpha, \beta \in \mathbb{T}$. If there is a sequence z_n in $\mathbb{D}$ such that $z_n \to \alpha$, $f(z_n) \to \beta$, and*

$$\lim_{n\to\infty} \frac{1-|f(z_n)|}{1-|z_n|} < \infty,$$

then

$$\angle \lim_{z\to\alpha} f(z) = \beta.$$

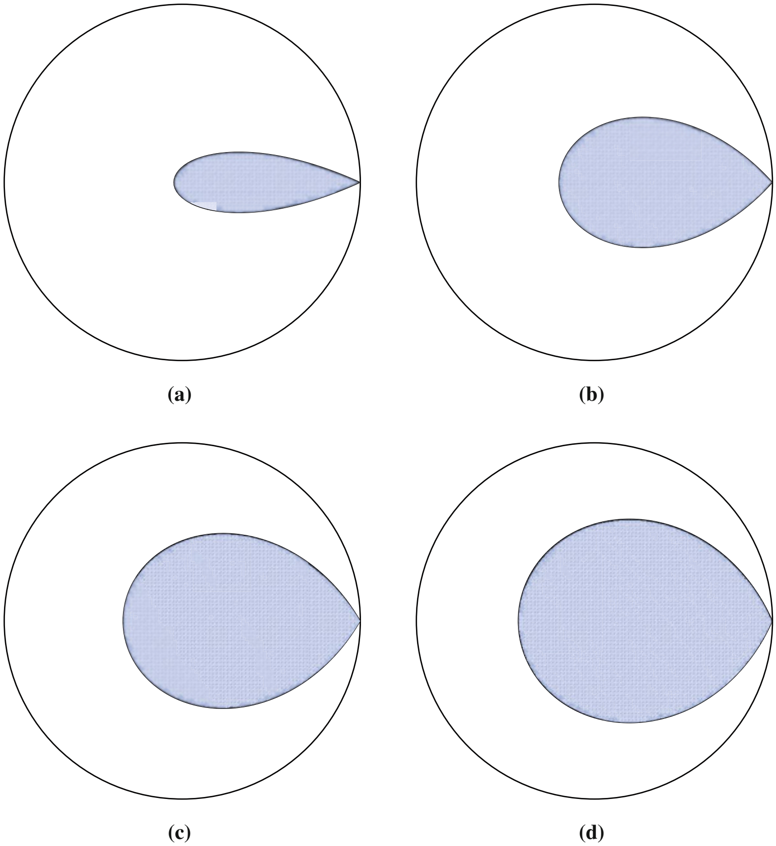


Fig. 1.1 Four Stolz domains anchored at 1. (**a**) $S_{1.1}(1)$. (**b**) $S_{1.5}(1)$. (**c**) $S_{2.0}(1)$. (**d**) $S_{2.5}(1)$

Proof Corollary 1.6.5 provides an $A > 0$ such that

$$\frac{|f(z)-\beta|^2}{1-|f(z)|^2} \leqslant A\frac{|z-\alpha|^2}{1-|z|^2}, \qquad z \in \mathbb{D}.$$

If $z \in S_C(\alpha)$, then

$$\begin{aligned} |f(z)-\beta|^2 &\leqslant \frac{|f(z)-\beta|^2}{1-|f(z)|^2} \\ &\leqslant A\frac{|z-\alpha|^2}{1-|z|^2} \\ &\leqslant AC\frac{|z-\alpha|}{1+|z|} \\ &\leqslant AC|z-\alpha|. \end{aligned}$$

Consequently,

$$\lim_{\substack{z\to\alpha \\ z\in S_C(\alpha)}} f(z) = \beta.$$

This holds for every Stolz domain anchored at α and thus $\angle \lim_{z\to\alpha} f(z) = \beta$. □

We can say a bit more. Under the hypotheses of the preceding corollary, one can conclude that f tends to β on certain domains that are tangential to $\mathbb{T}$ at α. However, Corollary 1.6.7 suffices for our applications.

1.7 Fixed Points

We say that $z_0 \in \mathbb{D}$ is a *fixed point* of $f \in \mathscr{S}$ if $f(z_0) = z_0$. For example, every point in $\mathbb{D}$ is a fixed point of the identity function $\mathrm{id}(z) = z$. On the other hand, a nonidentity rotation $\rho_\gamma(z) = \gamma z$, in which $\gamma \in \mathbb{T}\backslash\{1\}$, has only one fixed point in $\mathbb{D}$, namely 0. In this section, we investigate and classify the fixed points of automorphisms. We start by considering a more general problem.

Lemma 1.7.1 *If $f \in \mathscr{S}$ has two distinct fixed points in $\mathbb{D}$, then $f = \mathrm{id}$.*

Proof Suppose that $f \in \mathscr{S}$ has two distinct fixed points $\alpha, \beta \in \mathbb{D}$. Then

$$\left|\frac{f(\alpha) - f(\beta)}{1 - \overline{f(\beta)} f(\alpha)}\right| = \left|\frac{\alpha - \beta}{1 - \overline{\beta}\alpha}\right|,$$

so Theorem 1.4.1 says that $f \in \mathrm{Aut}(\mathbb{D})$. Theorem 1.2.5 provides $w \in \mathbb{D}$ and $\gamma \in \mathbb{T}$ such that

$$f(z) = \gamma \frac{w - z}{1 - \overline{w} z}. \tag{1.7.2}$$

The fixed points of f in $\mathbb{D}$ are the solutions to

$$z = \gamma \frac{w - z}{1 - \overline{w} z}$$

that belong to $\mathbb{D}$. Thus, $z \in \mathbb{D}$ is a fixed point of f if and only if

$$\overline{w} z^2 - (1 + \gamma) z + \gamma w = 0 \quad \text{and} \quad z \in \mathbb{D}. \tag{1.7.3}$$

There are three cases to consider.

- If $w \neq 0$, then the two solutions to (1.7.3) must be α and β. Thus,

$$(z - \alpha)(z - \beta) = z^2 - \frac{1 + \gamma}{\overline{w}} z + \gamma \frac{w}{\overline{w}} = 0$$

and hence

$$|\alpha||\beta| = \left|\gamma \frac{w}{\overline{w}}\right| = 1.$$

However, this last identity is impossible since $\alpha, \beta \in \mathbb{D}$.

- If $w = 0$ and $\gamma \neq -1$, then (1.7.3) reduces to a linear equation and hence has only one solution. This is a contradiction.
- If $w = 0$ and $\gamma = -1$, then every point in $\mathbb{D}$ is fixed; that is, $f = \mathrm{id}$.

This completes the proof. □

A variation of the preceding argument shows that any nonidentity Möbius transformation (1.3.8) has at most two fixed points in $\widehat{\mathbb{C}}$.

A closer look at the proof of Lemma 1.7.1 enables us to classify automorphisms based upon the number and location of their fixed points. Consider the automorphism (1.7.2), which is meromorphic on $\widehat{\mathbb{C}}$. Its fixed points in $\widehat{\mathbb{C}}$ are the solutions to (1.7.3). If $f \in \mathrm{Aut}(\mathbb{D})$ and $f \not\equiv \mathrm{id}$, then either f has exactly one fixed point inside $\mathbb{D}$ or two fixed points (possibly with repetition) on $\mathbb{T}$.

In light of the preceding discussion, we introduce the following definitions.

(a) An automorphism is *elliptic* if it has exactly one fixed point in $\mathbb{D}$.
(b) An automorphism is *hyperbolic* if it has two distinct fixed point on $\mathbb{T}$.
(c) An automorphism is *parabolic* if it has one repeated fixed point in $\mathbb{T}$. This case happens if and only if $w \neq 0$ and

$$(\gamma + 1)^2 = 4\gamma |w|^2.$$

The automorphism $f(z) = iz$ is elliptic; its fixed point in $\mathbb{D}$ is 0. An example of a hyperbolic automorphism is

$$f(z) = \frac{z - \frac{1}{2}}{1 - \frac{1}{2}z};$$

its fixed points are 1 and -1. Finally,

$$f(z) = i\,\frac{z - (\frac{1}{2} + \frac{i}{2})}{1 - \left(\frac{1}{2} - \frac{i}{2}\right)z}$$

is an example of a parabolic automorphism. Its only fixed point in the closed disk $\mathbb{D}^-$ is 1, which has multiplicity two. That is, the multiplicity of the zero of the rational function $f(z) - z$ at 1 is two.

1.8 Exercises

1.1 Prove Corollary 1.3.7 by direct computation.

1.2 Let $\Omega = \mathbb{R} \times (-\pi/2, \pi/2)$ and suppose that $f : \Omega \to \Omega$ is analytic.

(a) Show that $|f'(x)| \leqslant 1$ for all $x \in \mathbb{R}$.
(b) If $|f'(x_0)| = 1$ for some $x_0 \in \mathbb{R}$, show that $f(x_0) \in \mathbb{R}$. Find the general form of f.

Hint: Find appropriate conformal mappings $\phi : \mathbb{D} \to \Omega$ with $\phi(0) = x_0$ and $\psi : \Omega \to \mathbb{D}$ with $\psi(f(x_0)) = 0$ and then apply Lemma 1.1.1 to $\psi \circ f \circ \phi$.

1.3 Let $\Omega = \mathbb{R} \times (-\pi/2, \pi/2)$ and suppose that $f : \Omega \to \Omega$ is analytic.

(a) Show that

$$|f(x_2) - f(x_1)| \leqslant |x_2 - x_1|, \quad x_1, x_2 \in \mathbb{R}.$$

(b) Show that equality holds for a pair $x_1 \neq x_2$ if and only if

$$f(z) = z + c \quad \text{or} \quad f(z) = -z + c$$

for some constant $c \in \mathbb{R}$.

Hint: Note that

$$f(x_2) - f(x_1) = \int_{x_1}^{x_2} f'(x)\, dx$$

and then apply Exercise 1.2.

1.4 Fix $w \in (-1, 1)$. Show that τ_w maps $\mathbb{D}_- := \mathbb{D} \cap \{z : \operatorname{Im} z < 0\}$ bijectively onto $\mathbb{D}_+ := \mathbb{D} \cap \{z : \operatorname{Im} z > 0\}$, and vice versa.

1.5 Fix $w \in \mathbb{D}$. The function

$$g(z, w) : \mathbb{D} \to \mathbb{R} \cup \{+\infty\}$$

defined by

$$g(z, w) = \log \left| \frac{1 - \overline{w} z}{w - z} \right|$$

is the *Green's function* of $\mathbb{D}$ with singularity at w.

(a) Show that $g > 0$, $g(w, w) = +\infty$, and g is harmonic on $\mathbb{D} \setminus \{w\}$.
(b) Show that

$$\lim_{|z| \to 1} g(z, w) = 0.$$

Remark The function

$$z \mapsto g(z, w) + \log |w - z|$$

is a bounded harmonic function on $\mathbb{D}$. This reveals the logarithmic nature of the singularity of g at w.

1.6 Let $\mathbb{C}_+ := \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$. Show that $f : \mathbb{C}_+ \to \mathbb{C}_+$ is an automorphism of $\mathbb{C}_+$ if and only if

$$f(z) = \frac{az + b}{cz + d},$$

in which $a, b, c, d \in \mathbb{R}$ and $ad - bc > 0$.
Hint: Use the conformal mapping

$$z \mapsto \frac{z - i}{z + i},$$

which maps $\mathbb{C}_+$ onto $\mathbb{D}$, and then apply Theorem 1.2.5. State and prove corresponding results for $\mathbb{C}_- := \{z \in \mathbb{C} : \operatorname{Im} z < 0\}$.

1.7 Let $f \in \operatorname{Aut}(\mathbb{D})$ be represented as $f = \rho_\gamma \circ \tau_w$. Show that

$$\gamma = -\frac{f'(0)}{1 - |f(0)|^2} = -\frac{f'(0)}{|f'(0)|}.$$

Hint: By (1.3.2), $f'(0) = -\gamma(1 - |w|^2)$. Also note that $|w| = |f(0)|$ and $|\gamma| = 1$.

1.8 Show that if z_1 and z_2 are distinct points in $\mathbb{D}$, then there is an $f \in \operatorname{Aut}(\mathbb{D})$ so that $f(z_1) = 0$ and $0 < f(z_2) < 1$.
Hint: Consider $\gamma\tau_{z_1}$ for a suitable unimodular constant ζ.

1.9 Suppose that $\zeta_1, \zeta_2, w_1, w_2 \in \mathbb{T}$ are such that $\zeta_1 \neq \zeta_2$ and $w_1 \neq w_2$. Show that there is an $f \in \operatorname{Aut}(\mathbb{D})$ such that

$$f(\zeta_1) = w_1 \quad \text{and} \quad f(\zeta_2) = w_2.$$

Hint: First suppose that $\zeta_1 = 1$ and $\zeta_2 = -1$. Then appropriately compose two such functions.

1.10 Show that

$$(\rho_\gamma \circ \tau_w)^{-1} = \tau_w \circ \rho_{\overline{\gamma}} = \rho_{\overline{\gamma}} \circ \tau_{\gamma w}.$$

Hint: Use (1.2.3), (1.2.4), and Corollary 1.3.7.

1.11 Let $f \in \mathscr{S}$ and $z_0 \in \mathbb{D}$. Define

$$g(z) = f(w(z)), \quad z \in \mathbb{D},$$

in which $w(z) = \tau_{z_0}(z)$. Show that

$$|g'(z)|(1 - |z|^2) = |f'(w)|(1 - |w|^2).$$

Hint: Use (1.4.8).

1.12 Show that for all $f \in \mathscr{S}$,

$$\left|\frac{f(z) - f(w)}{z - w}\right|^2 \leqslant \frac{1 - |f(z)|^2}{1 - |z|^2} \cdot \frac{1 - |f(w)|^2}{1 - |w|^2}.$$

Hint: Use (1.6.1) and Theorem 1.4.1.

1.13 Let f be analytic on the disk $R\mathbb{D} = \{z \in \mathbb{C} : |z| < R\}$ and bounded there by M. Show that

$$\left|\frac{f(z) - f(w)}{z - w}\right| \leqslant \frac{2MR}{|R^2 - \overline{w}z|}, \qquad z, w \in R\mathbb{D}.$$

Hint: Consider $g(z) = f(Rz)/M$ for $z \in \mathbb{D}$, and apply Theorem 1.4.1.

1.14 Let $f : \mathbb{C}_+ \to \mathbb{C}_+$ be analytic. Show that

$$\left|\frac{f(z) - f(w)}{f(z) - \overline{f(w)}}\right| \leqslant \left|\frac{z - w}{z - \overline{w}}\right|, \qquad z, w \in \mathbb{C}_+.$$

Hint: Use the conformal mapping

$$z \mapsto \frac{z - w}{z - \overline{w}},$$

which maps $\mathbb{C}_+$ onto $\mathbb{D}$, and then apply Theorem 1.4.1.

1.15 Let $f : \mathbb{C}_+ \to \mathbb{C}_+$ be analytic. Show that

$$\left|\frac{f(z) - f(w)}{f(z) - \overline{f(w)}}\right| = \left|\frac{z - w}{z - \overline{w}}\right|, \qquad z, w \in \mathbb{C}_+,$$

if and only if

$$f(z) = \frac{az + b}{cz + d},$$

where $a, b, c, d \in \mathbb{R}$ and $ad - bc > 0$.
Hint: See Exercises 1.6 and 1.14.

1.16 Let $f : \mathbb{D} \to \mathbb{D}^-$ be analytic. Suppose that there is a sequence z_n in $\mathbb{D}$ such that $z_n \to 1$, $f(z_n) \to 1$, and

$$\lim_{n\to\infty} \frac{1 - |f(z_n)|}{1 - |z_n|} = 0.$$

Show that $f \equiv 1$. *Hint*: Use Julia's lemma.

1.17 Show that equality holds in Julia's inequality (1.6.4) if and only if $f \in \mathrm{Aut}(\mathbb{D})$.

Chapter 2
Elementary Hyperbolic Geometry

As a subset of the complex plane, the unit disk $\mathbb{D}$ inherits the standard Euclidean metric

$$d(z, w) := |z - w|.$$

However, there are other metrics on $\mathbb{D}$ that are more natural from the perspective of complex function theory. In this chapter we introduce the pseudohyperbolic and Poincaré metrics on $\mathbb{D}$. We also discuss the relationship between the curvature of a metric and the Schwarz Lemma (Lemma 1.1.1), along with hyperbolic geometry in the upper half-plane $\mathbb{C}_+$.

2.1 Pseudohyperbolic Metric

Definition 2.1.1 The *pseudohyperbolic metric* ρ on $\mathbb{D}$ is defined by

$$\varrho(z, w) := \left| \frac{z - w}{1 - \overline{w}z} \right|, \qquad z, w \in \mathbb{D}.$$

Since $|1 - \overline{w}z| < 2$ for all $z, w \in \mathbb{D}$, we have

$$\frac{1}{2} d(z, w) \leqslant \varrho(z, w). \tag{2.1.2}$$

For each compact subset K of $\mathbb{D}$, the expression $|1 - \overline{w}z|$ is bounded away from zero for $z, w \in K$. Consequently, there is a constant C_K such that

$$\varrho(z, w) \leqslant C_K d(z, w), \qquad z, w \in \mathbb{D}. \tag{2.1.3}$$

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_2

We will soon show that ϱ is a metric on $\mathbb{D}$ and that $(\mathbb{D}, \varrho)$ is a complete metric space. In contrast, the metric space $(\mathbb{D}, d)$ is not complete since the sequence $z_n = 1-\frac{1}{n}$ is Cauchy with respect to d but does not converge to a limit in $\mathbb{D}$. Consequently, it is more natural to endow $\mathbb{D}$ with the metric ϱ (or the closely related Poincaré metric) rather than the standard Euclidean metric.

The definition of ϱ ensures that $\varrho(z, w) = \varrho(w, z)$ and $0 \leqslant \varrho(z, w) < 1$ for all $z, w \in \mathbb{D}$, and that $\varrho(z, w) = 0$ if and only if $z = w$. To verify that ϱ satisfies the triangle inequality is more involved. We defer the proof until Sect. 2.2. For the moment, we assume that ϱ is a metric on $\mathbb{D}$.

The following restatement of the Schwarz–Pick Theorem (Theorem 1.4.1) says that any function in $\mathscr{S}$ (the Schur class) is a contraction on $\mathbb{D}$ with respect to the pseudohyperbolic metric ϱ.

Theorem 2.1.4 (Schwarz–Pick) *For $f \in \mathscr{S}$,*

$$\varrho(f(z), f(w)) \leqslant \varrho(z, w), \qquad z, w \in \mathbb{D}.$$

Moreover, equality holds for two distinct z, w if and only if $f \in \operatorname{Aut}(\mathbb{D})$.

For each $z_0 \in \mathbb{D}$ and $r \in (0, 1)$, the *pseudohyperbolic disk* with radius r centered at z_0 is

$$\Delta(z_0, r) := \{z \in \mathbb{D} : \varrho(z, z_0) < r\}.$$

Since $\varrho(z, z_0) = |\tau_{z_0}(z)|$ (recall (1.2.2)), it follows that $\Delta(z_0, r)$ is the inverse image of the Euclidean disk

$$D(0, r) := \{w \in \mathbb{C} : |w| < r\}$$

under the automorphism τ_{z_0}; that is,

$$\Delta(z_0, r) = \tau_{z_0}(D(0, r)), \tag{2.1.5}$$

since $\tau_{z_0}^{-1} = \tau_{z_0}$. To concretely describe $\Delta(z_0, r)$ requires the following result about Möbius transformations.

Lemma 2.1.6 *The image of a line or a circle under a Möbius transformation is a line or a circle.*

Proof Let

$$f(z) = \frac{az + b}{cz + d}, \tag{2.1.7}$$

in which $ad - bc \neq 0$. If $c = 0$, then f is a linear function and the desired result holds. If $c \neq 0$, then

$$f(z) = \frac{az + b}{cz + d} = \frac{\frac{a}{c}(cz + d) - \frac{ad}{c} + b}{cz + d} = \frac{(b - \frac{ad}{c})}{cz + d} + \frac{a}{c}$$

is a composition of linear functions and an inversion. Thus, it suffices to show that $z \mapsto \frac{1}{z}$ has the desired property. A circle in $\mathbb{R}^2$ is determined by an equation of the form

$$Ax + By + C(x^2 + y^2) = D, \tag{2.1.8}$$

in which $A, B, C, D \in \mathbb{R}$. Straight lines correspond to $C = 0$. If $z = x + iy$, then

$$\frac{1}{z} = \left(\frac{x}{x^2 + y^2}\right) + i\left(\frac{-y}{x^2 + y^2}\right) = u + iv.$$

Divide (2.1.8) by $x^2 + y^2$, write the result in terms of u and v, and obtain

$$Au - Bv - D(u^2 + v^2) = -C.$$

This is the equation of a circle if $D \neq 0$ or a line if $D = 0$. □

The preceding lemma asserts that the image of a Euclidean disk under a Möbius transformation is either another Euclidean disk or a half plane. Since $\tau_{z_0} \in \mathrm{Aut}(\mathbb{D})$, the inverse image under τ_{z_0} of any Euclidean disk contained in $\mathbb{D}$ is a Euclidean disk contained in $\mathbb{D}$. Consequently, (2.1.5) implies that the pseudohyperbolic disk $\Delta(z_0, \rho_0)$ is a Euclidean disk contained in $\mathbb{D}$. Can we be more specific?

By (2.1.5) we have

$$\Delta(z_0, r_0) = \left\{\frac{z_0 - re^{i\vartheta}}{1 - \overline{z_0}re^{i\vartheta}} : 0 \leqslant r < r_0,\ 0 \leqslant \vartheta \leqslant 2\pi\right\}. \tag{2.1.9}$$

Thus,

$$\partial\Delta(z_0, r_0) = \left\{\frac{z_0 - r_0e^{i\vartheta}}{1 - \overline{z_0}r_0e^{i\vartheta}} : 0 \leqslant \vartheta \leqslant 2\pi\right\}$$

is the circle that forms the boundary of $\Delta(z_0, r_0)$. The identity

$$\left|\frac{z_0 - r_0e^{i\vartheta}}{1 - \overline{z_0}r_0e^{i\vartheta}} - \frac{1 - |z_0|^2}{1 - r_0^2|z_0|^2}z_0\right| = \frac{1 - r_0^2}{1 - r_0^2|z_0|^2}r_0, \qquad \vartheta \in [0, 2\pi],$$

which can verified by direct computation (see Exercise 2.1), confirms that $\Delta(z_0, r_0)$ is precisely the Euclidean disk $D(c_0, R_0)$, in which

$$c_0 = \frac{1 - r_0^2}{1 - r_0^2|z_0|^2}z_0 \tag{2.1.10}$$

is the center of the corresponding Euclidean disk $D(c_0, R_0)$ and

$$R_0 = \frac{1 - |z_0|^2}{1 - r_0^2|z_0|^2}r_0 \tag{2.1.11}$$

is its radius. This is illustrated in Fig. 2.1.

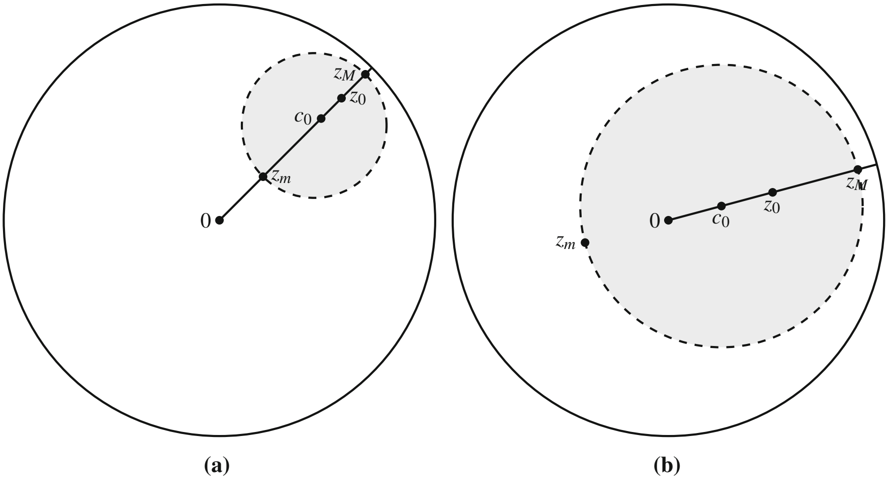


Fig. 2.1 Pseudohyperbolic disks. (**a**) $0 < r_0 \leqslant |z_0|$. (**b**) $|z_0| < r_0 < 1$

As (2.1.10) shows, c_0 lies on the line segment $[0, z_0]$ that joins 0 and z_0. Consequently, the point of maximum modulus on the boundary of $\Delta(z_0, r_0)$ is

$$z_M = c_0 + R_0 \frac{z_0}{|z_0|} = \frac{|z_0| + r_0}{1 + r_0|z_0|} \cdot \frac{z_0}{|z_0|}. \tag{2.1.12}$$

If $0 < r_0 \leqslant |z_0|$, then

$$z_m = c_0 - R_0 \frac{z_0}{|z_0|} = \frac{|z_0| - r_0}{1 - r_0|z_0|} \cdot \frac{z_0}{|z_0|} \tag{2.1.13}$$

is the point of minimum modulus on $\partial\Delta(z_0, r_0)$; otherwise $\Delta(z_0, r_0)$ contains 0. This is illustrated in Fig. 2.1b. The antipodal points z_m and z_M satisfy

$$1 - |z_M| = \frac{(1 - r_0)(1 - |z_0|)}{1 + r_0|z_0|} \leqslant 1 - |z| \leqslant \frac{(1 + r_0)(1 - |z_0|)}{1 - r_0|z_0|} = 1 - |z_m|$$

and hence

$$\frac{(1 - r_0)(1 - |z_0|)}{2} \leqslant 1 - |z| \leqslant 2 \frac{1 - |z_0|}{1 - r_0|z_0|} \tag{2.1.14}$$

for all $z \in \Delta(z_0, r_0)$. We leave it to the reader to obtain an appropriate estimate when $|z_0| < r_0$. In summary,

$$|z - c_0| \leqslant R_0 \quad \Longleftrightarrow \quad \left| \frac{z_0 - z}{1 - \overline{z}_0 z} \right| \leqslant r_0 \tag{2.1.15}$$

and

$$|z| \leqslant r_0 \quad \iff \quad \left|\frac{z_0 - z}{1 - \overline{z}_0 z} - c_0\right| \leqslant R_0. \tag{2.1.16}$$

The second equivalence follows from the first and (1.2.4). The parameters z_0, r_0, c_0, and R_0 are related by (2.1.10) and (2.1.11).

2.2 Generalized Triangle Inequality

In this section, we show that the pseudohyperbolic metric ϱ satisfies the triangle inequality. This is the final piece needed to verify that $(\mathbb{D}, \varrho)$ is a metric space. A little more work (see Theorem 2.2.4 below) shows that $(\mathbb{D}, \varrho)$ is complete.

Lemma 2.2.1 *If $z_0 \in \mathbb{D}$ and $r_0 \in (0, 1)$, then*

$$\frac{|z_0| - r_0}{1 - r_0|z_0|} \leqslant |z| \leqslant \frac{|z_0| + r_0}{1 + r_0|z_0|}, \qquad z \in \Delta(z_0, r_0)^-.$$

Proof Use (2.1.12) and (2.1.13). □

Corollary 2.2.2 *If $z_1, z_2 \in \mathbb{D}$, then*

$$\frac{z_1 - z_2}{1 - \overline{z}_1 z_2} \in \partial\Delta(z_1, |z_2|)$$

and

$$\frac{|z_1| - |z_2|}{1 - |z_1 z_2|} \leqslant \left|\frac{z_1 - z_2}{1 - \overline{z}_1 z_2}\right| \leqslant \frac{|z_1| + |z_2|}{1 + |z_1 z_2|}.$$

Proof By Theorem 2.1.4,

$$\varrho(z_1, \tau_{z_1}(z_2)) = \varrho(\tau_{z_1}(z_1), z_2) = \varrho(0, z_2) = |z_2|,$$

so $\tau_{z_1}(z_2) \in \partial\Delta(z_1, |z_2|)$. The desired inequalities follow from Lemma 2.2.1 with $z_0 = z_1$ and $r_0 = |z_2|$. □

We can apply Theorem 2.1.4 to the inequalities in Corollary 2.2.2 and obtain the following imposing inequalities that are not convenient to verify directly.

Theorem 2.2.3 *If $z_1, z_2, z_3 \in \mathbb{D}$, then*

$$\frac{\varrho(z_1, z_3) - \varrho(z_2, z_3)}{1 - \varrho(z_1, z_3)\varrho(z_2, z_3)} \leqslant \varrho(z_1, z_2) \leqslant \frac{\varrho(z_1, z_3) + \varrho(z_2, z_3)}{1 + \varrho(z_1, z_3)\varrho(z_2, z_3)}.$$

Proof We only prove the upper inequality since the proof of the lower inequality is similar. Indeed, the upper inequality holds for all $z_1, z_2, z_3 \in \mathbb{D}$ if and only if

$$\varrho(\tau_w(z_1), \tau_w(z_2)) \leqslant \frac{\varrho(\tau_w(z_1), \tau_w(z_3)) + \varrho(\tau_w(z_2), \tau_w(z_3))}{1 + \varrho(\tau_w(z_1), \tau_w(z_3))\varrho(\tau_w(z_2), \tau_w(z_3))}$$

for some $w \in \mathbb{D}$ and all $z_1, z_2, z_3 \in \mathbb{D}$. If $w = z_3$, then the preceding is equivalent to

$$\varrho(\tau_{z_3}(z_1), \tau_{z_3}(z_2)) \leqslant \frac{|\tau_{z_3}(z_1)| + |\tau_{z_3}(z_2)|}{1 + |\tau_{z_3}(z_1)\tau_{z_3}(z_2)|},$$

whose validity was previously established in Corollary 2.2.2. □

Theorem 2.2.3 shows that ϱ satisfies the triangle inequality

$$\varrho(z_1, z_2) \leqslant \varrho(z_1, z_3) + \varrho(z_3, z_2)$$

and hence $(\mathbb{D}, \varrho)$ is metric space. Unlike $(\mathbb{D}, d)$, in which d denotes the Euclidean metric, $(\mathbb{D}, \varrho)$ is complete. This emphasizes the importance of the pseudohyperbolic metric for function theory on the unit disk.

Theorem 2.2.4 *$(\mathbb{D}, \varrho)$ is a complete metric space.*

Proof If z_n is a ϱ-Cauchy sequence in $\mathbb{D}$, then it is ϱ-bounded (contained in a pseudo-hyperbolic disk). Lemma 2.2.1 shows that z_n is confined to a compact subset K of $\mathbb{D}$. Since (2.1.2) implies that z_n is Cauchy with respect to the Euclidean metric on K, the completeness of (K, d) implies that z_n converges with respect to the Euclidean metric to a limit $z \in K$. Then (2.1.3) shows that $z_n \to z$ with respect to ϱ. Thus, $(\mathbb{D}, \varrho)$ is complete. □

2.3 Poincaré Metric

For $z_1, z_2, z_3 \in \mathbb{D}$, Theorem 2.2.3 says that

$$\varrho(z_1, z_2) \leqslant \frac{\varrho(z_1, z_3) + \varrho(z_2, z_3)}{1 + \varrho(z_1, z_3)\varrho(z_2, z_3)}.$$

Rewrite this as

$$\frac{1 + \varrho(z_1, z_2)}{1 - \varrho(z_1, z_2)} \leqslant \frac{1 + \varrho(z_1, z_3)}{1 - \varrho(z_1, z_3)} \cdot \frac{1 + \varrho(z_2, z_3)}{1 - \varrho(z_2, z_3)}$$

and take logarithms to obtain

$$\log \frac{1 + \varrho(z_1, z_2)}{1 - \varrho(z_1, z_2)} \leqslant \log \frac{1 + \varrho(z_1, z_3)}{1 - \varrho(z_1, z_3)} + \log \frac{1 + \varrho(z_2, z_3)}{1 - \varrho(z_2, z_3)}. \tag{2.3.1}$$

This motivates the following definition.

Definition 2.3.2 The *Poincaré metric* on $\mathbb{D}$ is

$$\wp(z,w) = \log \frac{1+\varrho(z,w)}{1-\varrho(z,w)}, \qquad z,w \in \mathbb{D}.$$

Observe that $\wp(z,w) = \wp(w,z)$ and $\wp(z,w) \geqslant 0$ with equality if and only if $z = w$. The triangle inequality for $\wp$ is (2.3.1) and hence $(\mathbb{D}, \wp)$ is a metric space. Since

$$p(t) = \log\left(\frac{1+t}{1-t}\right)$$

is a strictly increasing function on $[0,1)$ and $\lim_{t\to 0^+} p(t) = 0$, it follows that the metric spaces $(\mathbb{D}, \varrho)$ and $(\mathbb{D}, \wp)$ have the same Cauchy sequences and the same convergent sequences. Consequently, Theorem 2.2.4 implies that $(\mathbb{D}, \wp)$ is a complete metric space. Theorem 2.1.4 can be rewritten as follows.

Theorem 2.3.3 (Schwarz–Pick) *If $f \in \mathscr{S}$, then*

$$\wp(f(z), f(w)) \leqslant \wp(z,w), \qquad z,w \in \mathbb{D}.$$

Equality holds for two distinct z, w if and only if $f \in \mathrm{Aut}(\mathbb{D})$.

Definition 2.3.4 The *hyperbolic length* of a simple piecewise C^1 curve Γ in $\mathbb{D}$ is

$$\ell(\Gamma) = \int_\Gamma \frac{2|dz|}{1-|z|^2}. \tag{2.3.5}$$

In the remainder of this book, the unmodified term "curve" always refers to a piecewise C^1 curve. What is the curve with the least hyperbolic length between two distinct points $z, w \in \mathbb{D}$? The Poincaré metric helps us answer this question.

Lemma 2.3.6 *Hyperbolic length is conformally invariant. That is, if Γ is a simple C^1 curve in $\mathbb{D}$ and $f \in \mathrm{Aut}(\mathbb{D})$, then*

$$\ell(f(\Gamma)) = \ell(\Gamma).$$

Proof Let $f \in \mathrm{Aut}(\mathbb{D})$ and let Γ be a curve in $\mathbb{D}$. Since (2.3.5) is rotationally invariant, it suffices to show that $\ell(\tau_{z_0}(\Gamma)) = \ell(\Gamma)$ for all $z_0 \in \mathbb{D}$. Let $w = \tau_{z_0}(z)$ and use (1.4.8) to obtain

$$\frac{|dw|}{1-|w|^2} = \frac{|\tau_{z_0}'(z)||dz|}{1-|\tau_{z_0}(z)|^2} = \frac{|dz|}{1-|z|^2}.$$

Then

$$\ell(\tau_{z_0}(\Gamma)) = \int_{\tau_{z_0}(\Gamma)} \frac{2|dw|}{1-|w|^2} = \int_\Gamma \frac{2|dz|}{1-|z|^2} = \ell(\Gamma). \qquad \square$$

The shortest distance between two points in Euclidean space is a straight line. What is the correct analogue in the hyperbolic setting? The following lemma concerns an instructive special case.

Lemma 2.3.7 *If* $0 < r < 1$*, then the real interval* $[0, r]$ *is the shortest hyperbolic curve between* 0 *and* r*. Moreover,*

$$\ell([0, r]) = \log \frac{1+r}{1-r};$$

that is, $\ell([0, r]) = \wp(0, r)$.

Proof First observe that

$$\begin{aligned}\ell([0, r]) &= \int_{[0,r]} \frac{2\,dz}{1-z^2} \\ &= \int_0^r \frac{2\,dt}{1-t^2} \\ &= \log \frac{1+r}{1-r}.\end{aligned}$$

Let Γ be a simple C^1 curve in $\mathbb{D}$ that starts at 0 and ends at r. Then the Cauchy integral formula, along with the fact that Γ is a simple, piecewise C^1 curve and hence homotopic to $[0, r]$, says that

$$\begin{aligned}\ell([0, r]) &= \left| \int_{[0,r]} \frac{2\,dz}{1-z^2} \right| \\ &= \left| \int_{\Gamma} \frac{2\,dz}{1-z^2} \right| \\ &\leqslant \int_{\Gamma} \frac{2\,|dz|}{1-|z|^2} \\ &= \ell(\Gamma).\end{aligned}$$

For the proof of uniqueness, see Exercise 2.3 and [128, Thm. 12.2.6a]. □

The following important theorem generalizes the preceding lemma.

Theorem 2.3.8 *The curve with the least hyperbolic length between two distinct points* $z_1, z_2 \in \mathbb{D}$ *is parametrized by*

$$\frac{z_1 - \tau_{z_1}(z_2)t}{1 - \overline{z_1}\tau_{z_1}(z_2)t}, \qquad t \in [0, 1].$$

Moreover, its hyperbolic length is $\wp(z_1, z_2)$.

Proof If $z_1, z_2 \in \mathbb{D}$ are distinct, then

$$f(z) = \frac{|\tau_{z_1}(z_2)|}{\tau_{z_1}(z_2)} \tau_{z_1}(z)$$

is an automorphism that satisfies $f(z_1) = 0$ and $0 < f(z_2) < 1$ (see Exercise 1.8). Thus, Lemmas 2.3.6 and 2.3.7 show that the curve with the shortest hyperbolic length between z_1 and z_2 is $f^{-1}([0, |\tau_{z_1}(z_2)|])$. Since $f^{-1}(z) = \tau_{z_1}(\overline{\gamma} z)$, where $\gamma = |\tau_{z_1}(z_2)|/\tau_{z_1}(z_2)$, a parametrization of the curve is

$$\begin{aligned}\Gamma(t) &= f^{-1}(t|\tau_{z_1}(z_2)|)\\ &= \tau_{z_1}(\overline{\gamma} t|\tau_{z_1}(z_2)|)\\ &= \tau_{z_1}(t\tau_{z_1}(z_2))\\ &= \frac{z_1 - \tau_{z_1}(z_2)t}{1 - \overline{z}_1\tau_{z_1}(z_2)t}\end{aligned}$$

for $t \in [0, 1]$. Theorem 2.3.3 and Lemma 2.3.7 imply that the length of this curve is

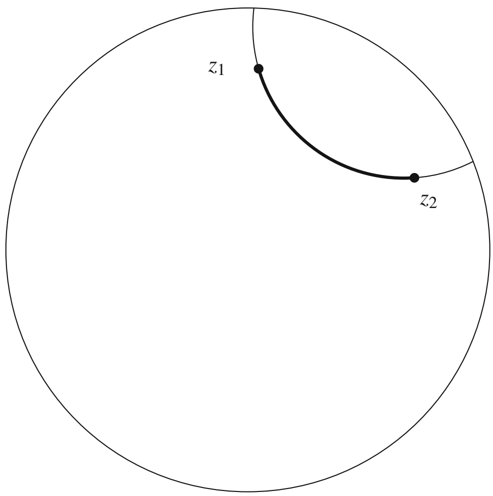
$$\wp(0, |\tau_{z_1}(z_2)|) = \wp(f^{-1}(0), f^{-1}(|\tau_{z_1}(z_2)|)) = \wp(z_1, z_2). \quad \square$$

Definition 2.3.9 The curve provided by Theorem 2.3.8 is the *hyperbolic line segment* between z_1 and z_2 (see Fig. 2.2). The (whole) *hyperbolic line* through two distinct points $z_1, z_2 \in \mathbb{D}$ is

$$w(t) = \frac{z_1 - (\frac{z_1 - z_2}{1 - \overline{z}_1 z_2})t}{1 - \overline{z}_1(\frac{z_1 - z_2}{1 - \overline{z}_1 z_2})t}, \qquad |t| < \left|\frac{1 - \overline{z}_1 z_2}{z_1 - z_2}\right|. \tag{2.3.10}$$

The restriction on t above guarantees that $w(t) \in \mathbb{D}$.

Fig. 2.2 The hyperbolic line segment between z_1 and z_2



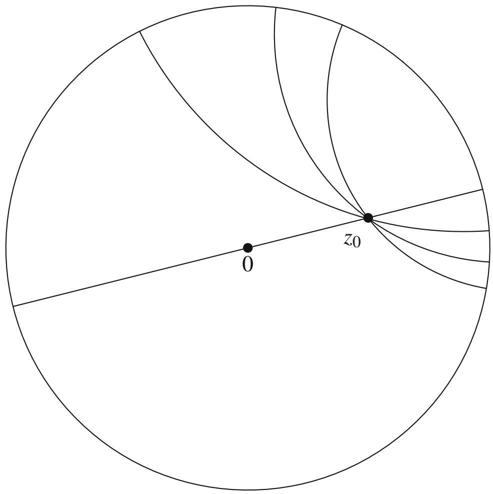


Fig. 2.3 Some hyperbolic lines that pass through z_0

The hyperbolic line segment between z_1 and z_2 is the image of the Euclidean line segment $[0, \tau_{z_1}(z_2)]$ under τ_{z_1}. Since the Euclidean line that passes through 0 and $\tau_{z_1}(z_2)$ is orthogonal to $\mathbb{T}$, Lemma 2.1.6 and the conformality of bijective analytic maps ensure that the hyperbolic line segment between z_1 and z_2 is either an arc of a circle orthogonal to $\mathbb{T}$, or a part of a diameter of $\mathbb{D}$ (see Fig. 2.3).

Corollary 2.3.11 *If $z_1, z_2, z_3 \in \mathbb{D}$ are distinct, then z_1, z_2, z_3 lie on the same hyperbolic line if and only if*

$$\Big(\frac{z_1 - z_2}{1 - \bar{z}_1 z_2}\Big)\Big/\Big(\frac{z_1 - z_3}{1 - \bar{z}_1 z_3}\Big) \in \mathbb{R}. \tag{2.3.12}$$

Proof Let z_1, z_2, and z_3 be distinct points in $\mathbb{D}$. Solve (2.3.10) for t and obtain

$$t = \Big(\frac{z_1 - z_2}{1 - \bar{z}_1 z_2}\Big)\Big/\Big(\frac{z_1 - w(t)}{1 - \bar{z}_1 w(t)}\Big).$$

Since z_3 lies on the hyperbolic line determined by z_1 and z_2 if and only if $z_3 = w(t)$ for some $t \in \mathbb{R}$, the desired result follows. □

2.4 Ahlfors's Version of the Schwarz's Lemma

Let μ be a positive, nonvanishing, twice continuously differentiable function on $\mathbb{D}$ and let Γ be a piecewise C^1 curve in $\mathbb{D}$. In differential geometry, μ is called a *metric* since it gives rise to a metric, in the metric-space sense, as follows. The length of Γ with respect to the metric μ is defined by

$$\ell_\mu(\Gamma) = \int_\Gamma \mu(z)\,|dz|.$$

For example, the Euclidean length of Γ is obtained when $\mu \equiv 1$. The Euclidean distance between two distinct points z and w in $\mathbb{D}$ is the length of the shortest piecewise C^1 curve between z and w. This curve is the straight line between z and w and its length is $|z - w|$. In a similar manner, the hyperbolic length (2.3.5) of a piecewise C^1 curve corresponds to the Poincaré metric

$$\mu(z) = \frac{2}{1 - |z|^2}. \tag{2.4.1}$$

If μ is a metric, in the geometric sense, on $\mathbb{D}$, then let $d_\mu(z, w)$ denote the length of the shortest piecewise C^1 curve between z and w. As a slight generalization of the concept above, we allow a metric μ to have isolated singularities. These points are usually the critical points of an analytic function or the pre-images of the singularities of another metric. This is further crystalized by the following construction.

If $f : \Omega_1 \to \Omega_2$ is analytic and μ is a metric on Ω_2, then its *pullback under* f is the metric $f^*\mu$ defined by

$$(f^*\mu)(z) = \mu(f(z))|f'(z)|, \qquad z \in \Omega_1.$$

If Γ is a piecewise C^1 curve in Ω_1, then a change of variables yields

$$\ell_{f^*\mu}(\Gamma) = \ell_\mu(f \circ \Gamma). \tag{2.4.2}$$

The situation becomes more interesting when $\Omega_1 = \Omega_2 = \Omega$. In this case, we can compare μ with its pullback $f^*\mu$ at each point of Ω. In light of (2.4.2), any such local relation between μ and $f^*\mu$ gives rise to a global relation between $\ell_\mu(f \circ \Gamma)$ and $\ell_\mu(\Gamma)$ for all curves Γ in Ω. Thus, we are led to a relation between $d_\mu(f(z), f(w))$ and $d_\mu(z, w)$, in which z, w are arbitrary points in Ω. We treat such possible phenomenon below.

If $f : \Omega \to \Omega$ is analytic and the metric μ is such that

$$(f^*\mu)(z) \leqslant \mu(z), \qquad z \in \Omega,$$

then

$$\ell_\mu(f \circ \Gamma) \leqslant \ell_\mu(\Gamma)$$

for all piecewise C^1 curves Γ. Thus,

$$d_\mu(f(z), f(w)) \leqslant d_\mu(z, w), \qquad z, w \in \Omega;$$

that is, f is a contraction from the metric space (Ω, d_μ) to itself.

We wish to focus on the Poincaré metric (2.4.1). With our new terminology, the second inequality in Theorem 1.4.1 becomes

$$(f^*\mu)(z) \leqslant \mu(z), \qquad z \in \mathbb{D}, \tag{2.4.3}$$

for any $f \in \mathscr{S}$. Theorem 2.3.3 is now a corollary of the approach above. In other words, any $f \in \mathscr{S}$ is a contraction in the hyperbolic setting. However, our goal is not only to restate the Schwarz–Pick Theorem in terms of this local metric. We wish to recast it in the language of differential geometry.

The *Laplace operator* is

$$\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

Note the difference in notation from (2.1.9). For the sake of simplicity, we are intentionally vague about the domain of definition of this differential operator, save that it operates on functions $f : \mathbb{D} \to \mathbb{C}$. In terms of the *Wirtinger derivatives*

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \quad \text{and} \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right), \tag{2.4.4}$$

one has

$$\triangle = 4\frac{\partial}{\partial z}\frac{\partial}{\partial \bar{z}} = 4\frac{\partial}{\partial \bar{z}}\frac{\partial}{\partial z}. \tag{2.4.5}$$

The *curvature* of a metric μ is

$$\kappa_\mu(z) = \frac{-\triangle \log \mu(z)}{\mu^2(z)}.$$

For example, the curvature of the Euclidean metric $\mu \equiv 1$ is identically 0. However, this is not the only metric on $\mathbb{D}$ that has constant curvature.

Lemma 2.4.6 *The curvature of the Poincaré metric* (2.4.1) *is identically equal to* -1.

Proof Since

$$\begin{aligned}\triangle \log \mu(z) &= \triangle \log \frac{2}{1-|z|^2} \\ &= 4\frac{\partial}{\partial z}\frac{\partial}{\partial \bar{z}} \log \frac{2}{1-z\bar{z}} \\ &= 4\frac{\partial}{\partial z}\frac{z}{1-z\bar{z}}\end{aligned}$$

$$= \frac{4}{(1 - \bar{z}z)^2}$$
$$= \frac{4}{(1 - |z|^2)^2}$$
$$= -\mu^2(z),$$

the desired result follows. □

Ahlfors realized that the Schwarz Lemma is a statement about curvature. His observation dramatically influenced the theory of functions.

Theorem 2.4.7 (Ahlfors [2]) *Let Ω be a domain in $\mathbb{C}$ endowed with a metric σ such that*

$$\kappa_\sigma(z) \leqslant -1, \quad z \in \Omega.$$

If $\mathbb{D}$ is endowed with the Poincaré metric μ and $f : \mathbb{D} \to \Omega$ is analytic, then

$$(f^*\sigma)(z) \leqslant \mu(z), \quad z \in \mathbb{D}.$$

Proof (Minda–Schober [103]) Fix $0 < r < 1$. On the disk $D(0, r)$ the metric

$$\mu_r(z) = \frac{2r}{r^2 - |z|^2}$$

is well defined. A similar calculation to the one used to prove Lemma 2.4.6 shows that

$$\kappa_{\mu_r}(z) = -1, \qquad z \in D(0, r).$$

Let

$$\Phi(z) = \frac{(f^*\sigma)(z)}{\mu_r(z)}, \qquad z \in D(0, r),$$

and observe that $\Phi \geqslant 0$. Since $f^*\sigma$ is continuous on $\mathbb{D}$, the function Φ is continuous on $D(0, r)$ and

$$\lim_{|z| \to r^-} \Phi(z) = 0.$$

Hence Φ attains its maximum M at some point of $D(0, r)^-$. Our goal is to show that $M \leqslant 1$.

If the maximum occurs on the boundary $\partial D(0, r)$ or at some point $z \in D(0, r)$ with $(f^*\sigma)(z) = 0$, then $M = \Phi \equiv 0$. Suppose that the maximum occurs at some point $z_0 \in D(0, r)$ and $(f^*\sigma)(z_0) > 0$. Since Φ is twice continuously differentiable at z_0,

$$\triangle \log \Phi(z_0) \leqslant 0.$$

According to our main hypothesis on the curvature of $f^*\sigma$,

$$\begin{aligned}\triangle \log \Phi(z_0) &= \triangle \log(f^*\sigma)(z_0) - \triangle \log \mu_r(z_0)\\ &= -\kappa_{f^*\sigma}(z_0)(f^*\sigma)^2(z_0) + \kappa_{\mu_r}(z_0)\mu_r^2(z_0)\\ &\geqslant (f^*\sigma)^2(z_0) - \mu_r^2(z_0).\end{aligned}$$

Therefore,

$$(f^*\sigma)^2(z_0) \leqslant \mu_r^2(z_0),$$

and so $\Phi(z_0) \leqslant 1$. In either case we have

$$(f^*\sigma)(z) \leqslant \rho_r(z), \qquad z \in D(0,r).$$

Let $r \to 1^-$ to complete the proof. □

When $\Omega = \mathbb{D}$ and $\sigma = \mu$ (the Poincaré metric), the preceding theorem is precisely the formulation (2.4.3) of the Schwarz–Pick Lemma.

2.5 Hyperbolic Geometry in $\mathbb{C}_+$

The preceding sections were concerned with hyperbolic geometry on the open unit disk $\mathbb{D}$. We can also discuss hyperbolic geometry in the upper half-plane $\mathbb{C}_+$. Instead of beginning anew with an independent approach, we use a conformal mapping between $\mathbb{D}$ and $\mathbb{C}_+$ to help establish our results.

The Möbius transformation

$$\varphi(z) = i\frac{1+z}{1-z} \tag{2.5.1}$$

is a bijection of $\mathbb{D}$ onto $\mathbb{C}_+$; see Fig. 2.4. Since

$$\varphi(1) = \infty, \qquad \varphi(-1) = 0, \qquad \varphi(i) = -1, \quad \text{and} \quad \varphi(-i) = 1,$$

Lemma 2.1.6 ensures that φ provides a bijection between $\mathbb{T}\backslash\{1\}$ and $\mathbb{R}$.

Since

$$\operatorname{Im}\varphi(z) = \frac{1-|z|^2}{|1-z|^2} \qquad \text{and} \qquad \varphi'(z) = \frac{2}{|1-z|^2},$$

it follows that

$$\frac{|\varphi'(z)|}{\operatorname{Im}\varphi(z)} = \frac{2}{1-|z|^2}, \qquad z \in \mathbb{D}. \tag{2.5.2}$$

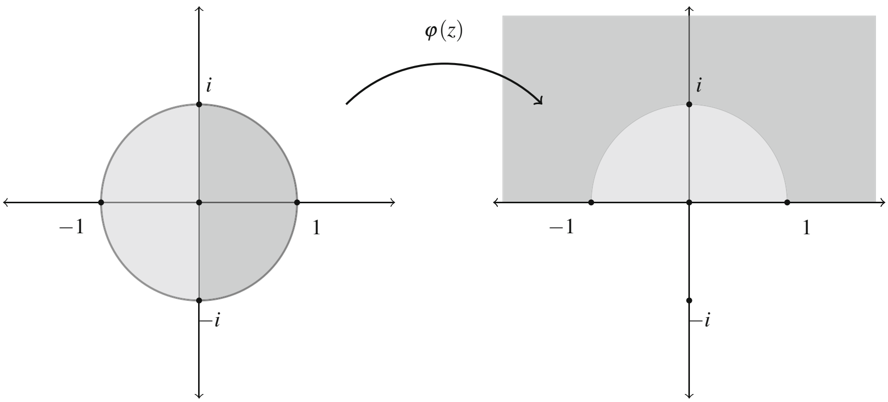


Fig. 2.4 The linear fractional transformation φ in (2.5.1)

Let Γ be a piecewise C^1 curve in $\mathbb{D}$ and let

$$\Upsilon = \varphi \circ \Gamma.$$

Then Υ is a piecewise C^1 curve in $\mathbb{C}_+$; indeed, $\varphi : \mathbb{D} \to \mathbb{C}_+$ provides a bijection between the family of piecewise C^1 curves in each domain. Consequently, (2.5.2) implies that

$$\frac{|\Upsilon'|}{\operatorname{Im} \Upsilon} = \frac{2|\Gamma'|}{1 - |\Gamma|^2}.$$

Since the hyperbolic length of the curve Γ in $\mathbb{D}$, as defined in (2.3.5), is

$$\ell(\Gamma) = \int_\Gamma \frac{2|dz|}{1 - |z|^2},$$

we define the hyperbolic length of the curve Υ in $\mathbb{C}_+$ by

$$\ell(\Upsilon) = \int_\Upsilon \frac{|dz|}{\operatorname{Im} z}. \tag{2.5.3}$$

2.6 Exercises

2.1 Verify by direct computation that $\Delta(z_0, \rho_0) = D(c_0, r_0)$, in which c_0 and r_0 are given by (2.1.10) and (2.1.11), respectively.

2.2 Show that the Poincaré metric $\wp$ satisfies

$$\wp(z, w) = 2 \tanh^{-1} \left| \frac{z - w}{1 - \overline{w} z} \right|, \qquad z, w \in \mathbb{D}.$$

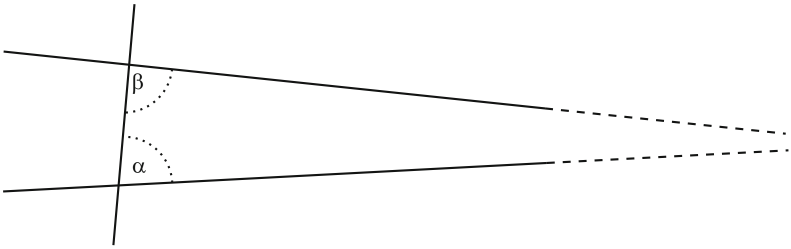


Fig. 2.5 Illustration of Euclid's fifth postulate. Here $\alpha + \beta < \pi$

Use the addition formula for the hyperbolic tangent function to derive the triangle inequality for the pseudohyperbolic metric.

2.3 Let $\gamma : [0, 1] \to \mathbb{C}$ be a C^1 path from 0 to $w \in \mathbb{D}\backslash\{0\}$ of minimal length.

(a) Why may we assume that γ does not vanish on $(0, 1]$?
(b) Write $\gamma(t) = r(t)e^{i\theta(t)}$, in which r, θ are C^1 functions on $[0, 1]$. Examine the integral that defines $\ell(\Gamma)$ and explain why θ is constant.
(c) Prove that a straight line is the shortest hyperbolic path from 0 to w.

See [128, Thm. 12.2.6a] for more on this.

2.4 Euclid's fifth postulate (the famous *Parallel Postulate*) asserts that "if two lines are drawn which intersect a third in such a way that the sum of the inner angles on one side is less than two right angles, then the two lines inevitably must intersect each other on that side if extended far enough." See Fig. 2.5 and Euclid's *Elements* [42].

(a) Prove that the Parallel Postulate is equivalent to Playfair's Axiom: "In a plane, given a line and a point not on it, at most one line parallel to the given line can be drawn through the point."
(b) In the Poincaré disk model of the hyperbolic plane, prove that given a line and a point not on it, there are infinitely many parallel lines to the given line that can be drawn through the point; see Fig. 2.6.

2.5 Suppose that $z_1, z_2, \ldots, z_n \in \mathbb{D}$ satisfy

$$\wp(z_i, z_j) = \begin{cases} 1 & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases} \tag{2.6.1}$$

(a) Show that such a configuration is impossible if $n \geqslant 5$.
Hint: Suppose that z_1, z_2, z_3, z_4, z_5 satisfy (2.6.1). Obtain circles Γ_1, Γ_2, both of the same hyperbolic radius, such that $z_2, z_3, z_4, z_5 \in \Gamma_1$ and $z_1, z_3, z_4, z_5 \in \Gamma_2$. Now examine $\Gamma_1 \cap \Gamma_2$.

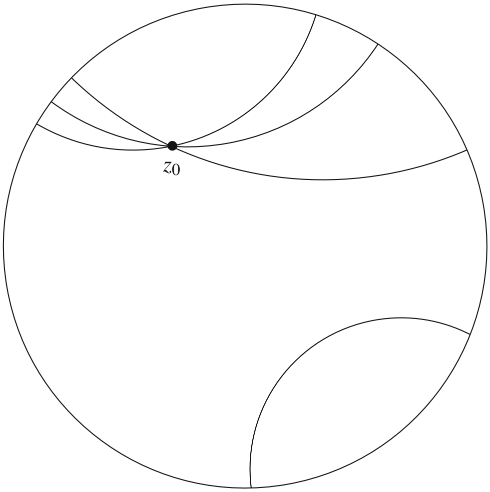


Fig. 2.6 Failure of the Parallel Postulate in the Poincaré disk model of hyperbolic geometry. Given a line that does not contain z_0, there are infinitely many hyperbolic lines through z_0 that are parallel to the given line

(b) Show that such a configuration is impossible if $n = 4$.
Hint: The proof is simple in the Euclidean case. The Poincaré model of the hyperbolic plane satisfies Hilbert's axioms [68, Sect. 39]. One can show that Propositions I.2-I.22 and I.24-I.28 of Euclid's *Elements* [42] can be obtained in the Poincaré model [68, Thm. 10.4]. Now proceed as in the Euclidean case.

2.6 In this exercise, we consider the Wirtinger differential operators $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial \overline{z}}$ defined in (2.4.4). For the sake of simplicity, assume that the functions involved are infinitely differentiable in the variables x and y.

(a) Verify that $\dfrac{\partial}{\partial z}$ and $\dfrac{\partial}{\partial \overline{z}}$ commute.
(b) Show that f is analytic if and only if $\dfrac{\partial}{\partial \overline{z}} f = 0$.
(c) Verify (2.4.5).

2.7 Let μ be a metric on $\mathbb{D}$ such that

$$\ell_\mu(f(\Gamma)) = \ell_\mu(\Gamma)$$

for all piecewise C^1 curves Γ and for all $f \in \mathrm{Aut}(\mathbb{D})$. Show that

$$\mu(z) = \frac{\mu(0)}{1 - |z|^2}, \qquad z \in \mathbb{D}.$$

This is the converse of Lemma 2.3.6.
Hint: Study the local behavior of $f = \tau_{z_0}$ at the origin.

2.8 The *spherical metric* is defined by

$$\sigma(z) = \frac{2}{1+|z|^2}, \qquad z \in \mathbb{D}.$$

Show that

$$\kappa_\sigma(z) = 1, \quad z \in \mathbb{D}.$$

Hint: Use (2.4.5).

Chapter 3
Finite Blaschke Products: The Basics

3.1 Finite Blaschke Products

Definition 3.1.1 For a finite sequence $z_1, z_2, \ldots, z_n$ in $\mathbb{D}$ and $\gamma \in \mathbb{T}$, the function

$$B(z) = \gamma \prod_{k=1}^{n} \frac{z - z_k}{1 - \overline{z_k} z} \tag{3.1.2}$$

is a *finite Blaschke product*.

In the preceding, we allow repetition of the z_j. For example, if

$$z_1 = z_2 = \cdots = z_n = 0 \quad \text{and} \quad \gamma = 1,$$

then $B(z) = z^n$.

The function B is a unimodular constant (constant of modulus one) times a product of the automorphisms $\tau_{z_1}, \tau_{z_2}, \ldots, \tau_{z_n}$ defined in (1.2.2). The finite Blaschke product B is a rational function with zeros at the z_j, and nowhere else. It has a meromorphic extension to $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ with poles at

$$1/\overline{z_1}, \quad 1/\overline{z_2}, \ldots, \quad 1/\overline{z_n},$$

all of which lie in

$$\mathbb{D}_e := \widehat{\mathbb{C}} \backslash \mathbb{D}^-,$$

the *extended exterior disk*. The use of the *extended* exterior disk is important here since $1/z_j = \infty$ when $z_j = 0$. There is also a notion of an infinite Blaschke product which we will discuss in the endnotes of this chapter.

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_3

Since a nonconstant finite Blaschke product B is a product of disk automorphisms, it belongs to the Schur class $\mathscr{S}$ (Definition 1.0.2) and satisfies

$$|B(z)| < 1, \qquad z \in \mathbb{D}, \tag{3.1.3}$$

$$|B(\zeta)| = 1, \qquad \zeta \in \mathbb{T}, \tag{3.1.4}$$

$$|B(z)| > 1, \qquad z \in \widehat{\mathbb{C}} \backslash \mathbb{D}. \tag{3.1.5}$$

For $\xi \in \mathbb{T}$, observe that $\xi = 1/\overline{\xi}$. By (3.1.4) we have

$$B(\zeta) = \frac{1}{\overline{B(\zeta)}} = \frac{1}{\overline{B(1/\overline{\zeta})}}, \qquad \zeta \in \mathbb{T}.$$

Since $1/\overline{B(1/\overline{z})}$ is meromorphic on $\widehat{\mathbb{C}}$ and agrees with $B(z)$ on $\mathbb{T}$, the two functions are identical, that is,

$$B(z) = \frac{1}{\overline{B(1/\overline{z})}}, \qquad z \in \widehat{\mathbb{C}}. \tag{3.1.6}$$

The preceding equality can also be obtained from (3.1.2) by a direct computation; see Exercise 3.4.

3.2 Uniqueness and Nonuniqueness

The alert reader might question why the unimodular constant γ is included in the definition of a Blaschke product. In the next several sections, we will characterize the finite Blaschke products among the functions in the Schur class and γ will play a role. On the rare occasions when we need to distinguish the finite Blaschke products where $\gamma = 1$, we will use the term *monic finite Blaschke product* for which we have the following uniqueness theorem.

Theorem 3.2.1 (Horowitz–Rubel [80]) *Suppose that*

$$B_1(z) = \prod_{k=1}^{n} \frac{z - z_k}{1 - \overline{z_k} z} \quad \textit{and} \quad B_2(z) = \prod_{k=1}^{n} \frac{z - \lambda_k}{1 - \overline{\lambda_k} z}$$

are two monic finite Blaschke products of degree n and $B_1(w_k) = B_2(w_k)$ for n distinct points $w_1, w_2 \ldots, w_n$ in $\mathbb{D}$. Then $B_1(z) = B_2(z)$ for all $z \in \mathbb{D}$.

We follow the original proof from [80] which requires the following lemma.

Lemma 3.2.2 *Suppose that*

$$B_1(z) = \prod_{k=1}^{n} \frac{z - z_k}{1 - \overline{z_k} z} \quad \textit{and} \quad B_2(z) = \prod_{k=1}^{n} \frac{z - \lambda_k}{1 - \overline{\lambda_k} z}$$

are two monic finite Blaschke products of degree n and $B_1(w_k) = B_2(w_k)$ *for n distinct points* $w_1, w_2 \ldots, w_n$ *in* $\mathbb{D}$. *Then* $B_1(\xi) = B_2(\xi)$ *for some* $\xi \in \mathbb{T}$.

Proof By cross multiplication of terms, note that

$$\prod_{k=1}^{n} \frac{z - z_k}{1 - \overline{z_k} z} = \prod_{k=1}^{n} \frac{z - \lambda_k}{1 - \overline{\lambda_k} z}$$

precisely when

$$\prod_{k=1}^{n} \Big(\frac{z - z_k}{z - \lambda_k}\Big) \Big/ \Big(\frac{1 - \overline{z_k} z}{1 - \overline{\lambda_k} z}\Big) = 1.$$

Moreover, when $|z| = 1$ the identity above is equivalent to

$$\prod_{k=1}^{n} \Big(\frac{z - z_k}{z - \lambda_k}\Big) \Big/ \overline{\Big(\frac{z - z_k}{z - \lambda_k}\Big)} = 1.$$

Since

$$w/\overline{w} = e^{2i \arg w}$$

for any $w \in \mathbb{C}\backslash\{0\}$, we will be done if we can show that

$$\arg \prod_{k=1}^{n} \frac{\xi - z_k}{\xi - \lambda_k} = \pi m$$

for some $\xi \in \mathbb{T}$ and some integer m.

To do this, define

$$F(z) = \prod_{k=1}^{n} \frac{z - z_k}{z - \lambda_k} \tag{3.2.3}$$

and observe that F is analytic and zero free on $|z| > 1 - \delta$ for some $\delta > 0$ and that $F(\infty) = 1$. Hence there is an analytic branch of $H(z) = \log F(z)$ for $|z| > 1 - \delta$. But since $F(\infty) = 1$, we may choose $H(\infty) = 0$.

Define

$$h(z) = \operatorname{Im}(H(1/z)), \quad |z| < \frac{1}{1 - \delta}$$

and note that h is harmonic on an open neighborhood of $\mathbb{D}^-$. By the Mean Value Property for harmonic functions,

$$0 = H(\infty) = h(0) = \int_0^{2\pi} h(e^{i\theta}) \frac{d\theta}{2\pi}.$$

Since h is continuous and real valued on $\mathbb{T}$, it follows that $h(\xi) = 0$ for some $\xi \in \mathbb{T}$. However, since $\operatorname{Im} H$ is a branch of $\arg F$ and any two branches of $\arg F$ differ by a constant integer multiple of 2π, the lemma is proved. □

Proof (of Theorem 3.2.1) Let R be the rational function defined by

$$R(z) = \frac{B_1(z)}{B_2(z)}$$

and note that since B_1 and B_2 are finite Blaschke products, $|R(\zeta)| = 1$ for all $\zeta \in \mathbb{T}$ by (3.1.4). By the same argument use to prove (3.1.6), we see that

$$R(z)\overline{R(1/\overline{z})} = 1, \quad z \in \widehat{\mathbb{C}}.$$

By hypothesis,

$$R(w_1) = R(w_2) = \cdots = R(w_n) = 1.$$

The previous identity also says that

$$R(1/w_1) = R(1/w_2) = \cdots = R(1/w_n) = 1.$$

Since $|w_j| < 1$, we see that $|1/w_j| > 1$. Furthermore, by the previous lemma there is a $\xi \in \mathbb{T}$ for which $R(\xi) = 1$. Putting this all together, we have $2n + 1$ points

$$w_1, w_2, \ldots, w_n, \quad 1/w_1, 1/w_2, \ldots, 1/w_n, \quad \xi$$

that are mapped to 1 by R. Since the degree of the rational function R is $2n$, we see that $R \equiv 1$ and thus $B_1 \equiv B_2$. □

The fact that B_1 and B_2 are *monic* is important and was used in (3.2.3) to get that $F(\infty) = 1$. If we do not assume that B_1 and B_2 are monic, the conclusion of Theorem 3.2.1 is not always true. Indeed, suppose that

$$z_1, z_2, \ldots, z_n \in \mathbb{D}\backslash\{0\}$$

and define

$$B_1(z) = i \prod_{k=1}^{n} \frac{z - z_k}{1 - \overline{z_k} z} \quad \text{and} \quad B_2(z) = \frac{z - i z_1}{1 + i \overline{z_1} z} \prod_{k=2}^{n} \frac{z - z_k}{1 - \overline{z_k} z}.$$

Then B_1 and B_2 are finite Blaschke products of degree n with $B_1(z_k) = B_2(z_k)$ for all $k = 2, 3, \ldots, n$ and $B_1(0) = B_2(0)$. However, $B_1(w) \neq B_2(w)$ for some $w \in \mathbb{D}$. Indeed, if this were not the case, then

$$1 = \frac{B_1(z)}{B_2(z)} = i \frac{\tau_{z_1}(z)}{\tau_{i z_1}(z)}, \quad z \in \mathbb{D}.$$

The preceding says that $\tau_{z_1} = i \tau_{i z_1}$, and hence, using the fact that $\tau_{i z_1} \circ \tau_{i z_1} = \mathrm{id}$ (see (1.2.4)), we get $(\tau_{z_1} \circ \tau_{i z_1})(z) = i$ for all $z \in \mathbb{D}$. This contradicts the fact that $\tau_{z_1} \circ \tau_{i z_1} \in \operatorname{Aut}(\mathbb{D})$.

3.3 Finite Blaschke Products as Rational Functions

Two polynomials are *relatively prime* if they have no nonconstant common factor. If P and Q are relatively prime polynomials and Q is not identically zero, then the *degree* of the rational function $f = P/Q$ is

$$\deg f := \max\{\deg P, \deg Q\},$$

where $\deg P$ and $\deg Q$ are the degrees of P and Q, respectively.

For a finite Blaschke product B, rewrite (3.1.2) to obtain

$$B(z) = \gamma \frac{\prod_{k=1}^{n}(z - z_k)}{\prod_{k=1}^{n}(1 - \overline{z_k} z)}, \tag{3.3.1}$$

which confirms that B is a rational function of degree n. We regard a unimodular constant function as a finite Blaschke product of degree 0. The following theorem shows that the numerator and denominator in (3.3.1) are closely related.

Theorem 3.3.2 *A rational function of degree n is a finite Blaschke product of degree n if and only if it is of the form*

$$\frac{z^n \overline{P(1/\overline{z})}}{P(z)} = \frac{\overline{\alpha}_n + \overline{\alpha}_{n-1} z + \cdots + \overline{\alpha}_0 z^n}{\alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n}, \tag{3.3.3}$$

in which $\alpha_0 \neq 0$ and the numerator has all of its zeros in $\mathbb{D}$.

Proof Suppose that B is a finite Blaschke product and write it in the form (3.3.1). Let $\gamma = e^{it_0}$ in (3.3.1) and define

$$P(z) = e^{-it_0/2} \prod_{k=1}^{n} (1 - \overline{z_k} z), \tag{3.3.4}$$

which is a polynomial of degree at most n. Since

$$\begin{aligned} z^n \overline{P(1/\overline{z})} &= z^n e^{it_0/2} \prod_{k=1}^{n} \overline{(1 - \overline{z_k}/\overline{z})} \\ &= z^n e^{it_0/2} \prod_{k=1}^{n} (1 - z_k/z) \\ &= e^{it_0/2} \prod_{k=1}^{n} (z - z_k), \end{aligned} \tag{3.3.5}$$

we obtain

$$\begin{aligned}\frac{z^n\overline{P(1/\overline{z})}}{P(z)} &= \frac{e^{it_0/2}\prod_{k=1}^{n}(z-z_k)}{e^{-it_0/2}\prod_{k=1}^{n}(1-\overline{z_k}z)}\\ &= \gamma\prod_{k=1}^{n}\frac{z-z_k}{1-\overline{z_k}z}\\ &= B(z) \end{aligned}\tag{3.3.6}$$

by (3.3.1). The numerator

$$z^n\overline{P(1/\overline{z})} = e^{it_0/2}\prod_{k=1}^{n}(z-z_k)$$

has all of its zeros in $\mathbb{D}$. It is a polynomial of degree n, so $\alpha_0 \neq 0$ in (3.3.3).

Conversely, suppose that f is a rational function of the form (3.3.3), in which $\alpha_0 \neq 0$ and the numerator has all of its zeros in $\mathbb{D}$. By scaling the numerator and denominator in (3.3.3) by a real constant factor, we may assume that $\alpha_0 \in \mathbb{T}$. Then the numerator $z^n\overline{P(1/\overline{z})}$ is of the form (3.3.5), in which $z_1, z_2, \ldots, z_n \in \mathbb{D}$. Consequently, P enjoys a factorization of the form (3.3.4) and

$$\frac{z^n\overline{P(1/\overline{z})}}{P(z)}$$

is of the form (3.3.1), so it is a finite Blaschke product. □

Fix $n > 0$. If P is a polynomial of degree at most n, then let $P^{\#_n}$ be the polynomial

$$P^{\#_n}(z) = z^n\overline{P(1/\overline{z})},$$

which has degree at most n. To be more specific,

$$(\alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n)^{\#_n} = \overline{\alpha}_n + \overline{\alpha}_{n-1}z + \cdots + \overline{\alpha}_0 z^n.$$

That is, $P^{\#_n}$ is obtained from P by reversing the coefficients and conjugating them. Observe that

$$(1)^{\#_n} = z^n \quad \text{and} \quad (z^n)^{\#_n} = 1.$$

We usually write # without reference to n since the value of n is determined by context. Observe that

$$(P^{\#})^{\#} = P \tag{3.3.7}$$

and the zeros of $P^{\#}$ are

$$1/\overline{z_1}, \quad 1/\overline{z_2}, \ldots, \quad 1/\overline{z_n},$$

in which $z_1, z_2, \ldots, z_n$ denote the zeros of P. Moreover,

$$(QR)^{\#_{j+k}} = Q^{\#_j} R^{\#_k} \tag{3.3.8}$$

whenever Q, R are nonconstant polynomials and j, k are positive integers.

The proof of Theorem 3.3.2 shows that a rational function is a finite Blaschke product if and only if it is of the form

$$P/P^{\#}, \tag{3.3.9}$$

in which P is a unimodular scalar multiple of

$$(z - z_1)(z - z_2)\cdots(z - z_n),$$

and $z_1, z_2, \ldots, z_n \in \mathbb{D}$. This result has the following generalization.

Corollary 3.3.10 *A rational function of degree n is a quotient of two finite Blaschke products whose degrees sum to at most n if and only if it is of the form*

$$\frac{P}{P^{\#}},$$

in which P is a polynomial with no zeros on $\mathbb{T}$.

Proof Suppose that P is a nonconstant polynomial with no zeros on $\mathbb{T}$. Write $P = QR^{\#}$, in which Q is a polynomial with all of its zeros in $\mathbb{D}$ and $R^{\#}$ is a polynomial with all of its zeros in $\mathbb{D}_e$. That is, $R = (R^{\#})^{\#}$ is a polynomial with all of its zeros in $\mathbb{D}$. Then (3.3.8) ensures that

$$f = \frac{P}{P^{\#}} = \frac{QR^{\#}}{(QR^{\#})^{\#}} = \frac{QR^{\#}}{Q^{\#}R} = \frac{Q}{Q^{\#}} \Big/ \frac{R}{R^{\#}}, \tag{3.3.11}$$

which is a quotient of two finite Blaschke products by Theorem 3.3.2. If Q and R have any zeros in common, then cancellation occurs in (3.3.11). After this, the degrees of the resulting finite Blaschke products sum to at most n.

Conversely, if f is a quotient of two finite Blaschke products whose degrees sum to at most n, then by Theorem 3.3.2 we may write

$$f = \frac{Q}{Q^{\#}} \Big/ \frac{R}{R^{\#}},$$

in which Q and R are polynomials with all of their zeros in $\mathbb{D}$ and so that

$$\deg Q + \deg R^{\#} \leqslant n.$$

Note that $P = QR^{\#}$ is a polynomial with no zeros on $\mathbb{T}$ and that $f = P/P^{\#}$. □

Corollary 3.3.10 implies that

$$|P(\zeta)| = |P^{\#}(\zeta)| \tag{3.3.12}$$

for all polynomials P and $\zeta \in \mathbb{T}$. See Exercise 3.7 for alternate proofs of Theorem 3.3.2 and Corollary 3.3.10.

3.4 Finite Blaschke Products as n-to-1 Functions

Let B be a finite Blaschke product (3.1.2) of degree n and consider the equation

$$B(z) = w, \tag{3.4.1}$$

in which $w \in \widehat{\mathbb{C}}$. Write $B = P/P^{\#}$ as in (3.3.9), where P is a polynomial of degree n. Then (3.4.1) becomes

$$P(z) - wP^{\#}(z) = 0, \tag{3.4.2}$$

in which $P - wP^{\#}$ is a polynomial of degree at most n. Consequently, for each $w \in \widehat{\mathbb{C}}$, (3.4.1) has n solutions in $\widehat{\mathbb{C}}$, repeated according to their multiplicity. For example, if $w = 0$, then the solutions are precisely the zeros $z_1, z_2, \ldots, z_n$ of B; that is, the zeros of P. As another example, if $w = \infty$, then the solutions are precisely $1/\overline{z_1}, 1/\overline{z_2}, \ldots, 1/\overline{z_n}$, which are the poles of B.

We now show that B has constant valence on each of the disk $\mathbb{D}$, the extended exterior disk $\mathbb{D}_e$, and the unit circle $\mathbb{T}$ (see Fig. 3.1). We get started on $\mathbb{T}$ with the following lemma.

Lemma 3.4.3 *If B is a finite Blaschke product, then $B'(\zeta) \neq 0$ for all $\zeta \in \mathbb{T}$.*

Proof To prove this result, we require the *logarithmic derivative* of a product

$$f = f_1 f_2 \cdots f_n$$

of meromorphic functions $f_1, f_2, \ldots, f_n$:

$$\frac{f'}{f} = \frac{f_1'}{f_1} + \frac{f_2'}{f_2} + \cdots + \frac{f_n'}{f_n}. \tag{3.4.4}$$

The logarithmic derivative of a finite Blaschke product

$$B(z) = \gamma \prod_{k=1}^{n} \frac{z - z_k}{1 - \overline{z_k} z}, \qquad \gamma \in \mathbb{T}, \tag{3.4.5}$$

is

$$\begin{aligned} \frac{B'(z)}{B(z)} &= \sum_{k=1}^{n} \left(\frac{z - z_k}{1 - \overline{z_k} z} \right)' \cdot \frac{1 - \overline{z_k} z}{z - z_k} \\ &= \sum_{k=1}^{n} \frac{(1 - \overline{z_k} z) - (z - z_k)(-\overline{z_k})}{(1 - \overline{z_k} z)^2} \cdot \frac{1 - \overline{z_k} z}{z - z_k} \\ &= \sum_{k=1}^{n} \frac{1 - |z_k|^2}{(1 - \overline{z_k} z)(z - z_k)}. \end{aligned} \tag{3.4.6}$$

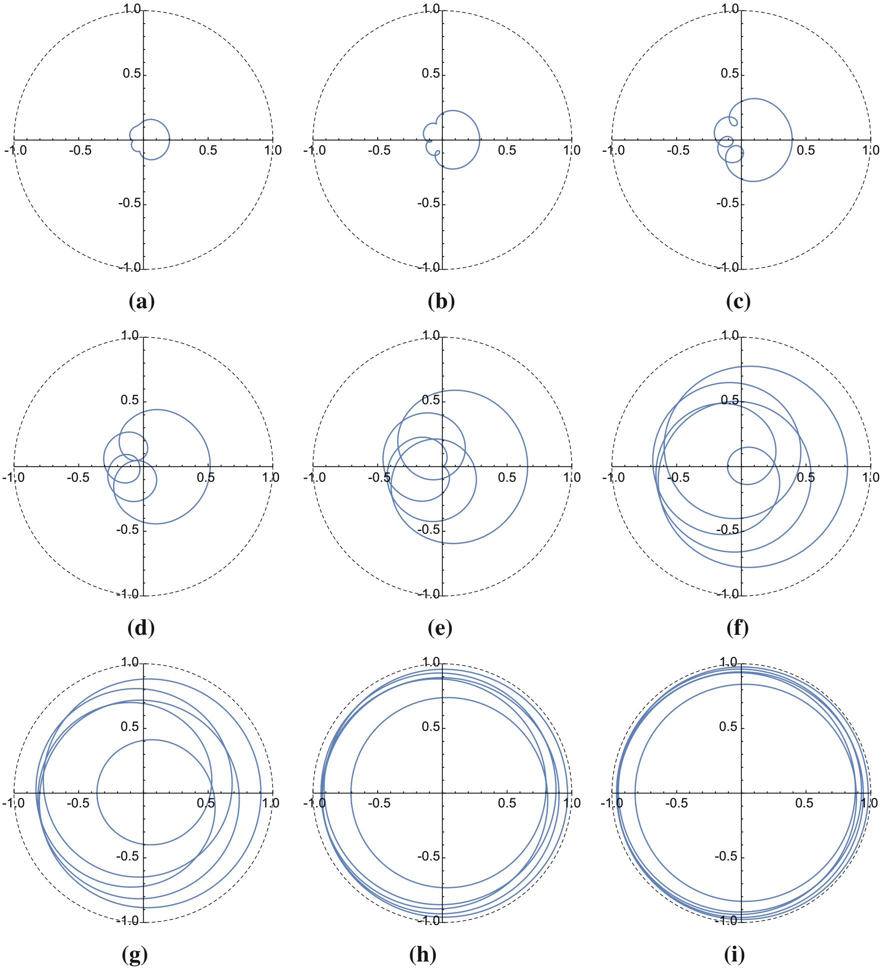


Fig. 3.1 Images of the circle $|z| = r$ under the finite Blaschke product B from (3.1.2) with $z_1 = 0$, $z_2 = z_3 = \frac{4}{5}$, $z_4 = \frac{2i}{3}$, and $z_5 = \frac{-3i}{4}$. All five zeros of B have modulus at most 0.8; hence the image of $|z| = r$ for $r > 0.8$ winds around the origin five times. Thus, B is a 5-to-1 map from $\mathbb{D}$ onto itself. (**a**) $r = 0.4$. (**b**) $r = 0.5$. (**c**) $r = 0.6$. (**d**) $r = 0.7$. (**e**) $r = 0.8$. (**f**) $r = 0.9$. (**g**) $r = 0.95$. (**h**) $r = 0.9825$. (**i**) $r = 0.99$

If $\zeta \in \mathbb{T}$, then

$$\zeta \frac{B'(\zeta)}{B(\zeta)} = \sum_{k=1}^{n} \frac{1 - |z_k|^2}{\overline{\zeta}(1 - \overline{z_k}\zeta)(\zeta - z_k)} = \sum_{k=1}^{n} \frac{1 - |z_k|^2}{|\zeta - z_k|^2}. \tag{3.4.7}$$

Since $B(\zeta)$ and ζ are both unimodular, it follows that

$$|B'(\zeta)| = \sum_{k=1}^{n} \frac{1-|z_k|^2}{|\zeta - z_k|^2} > 0. \quad \square \tag{3.4.8}$$

The identity in (3.4.8) ensures that a finite Blaschke product cannot assume any (necessarily unimodular) values on $\mathbb{T}$ with multiplicity greater than one.

This next result says that the argument of a finite Blaschke product is always increasing. Our proof follows [30].

Corollary 3.4.9 *For a finite Blaschke product B,*

$$\frac{d}{dt} \arg B(e^{it}) = |B'(e^{it})|.$$

Proof The calculation from (3.4.6) shows that

$$\frac{B'(z)}{B(z)} = \sum_{k=1}^{n} \frac{1-|z_k|^2}{(1-\overline{z_k}z)(z-z_k)}.$$

Writing $B(e^{it}) = e^{i\psi(t)}$, where $\psi(t)$ is real valued and $\psi(2\pi) - \psi(0)$ is a multiple of 2π, (3.4.7) says that

$$\psi'(t) = e^{it}\frac{B'(e^{it})}{B(e^{it})} = \sum_{k=1}^{n} \frac{1-|z_k|^2}{|e^{it}-z_k|^2}.$$

By (3.4.8), this last quantity equals $|B'(e^{it})|$. $\square$

Theorem 3.4.10 *Let B be a finite Blaschke product of degree n. For each $w \in \widehat{\mathbb{C}}$, the equation $B(z) = w$ has exactly n solutions in $\widehat{\mathbb{C}}$, counted according to multiplicity.*

(a) If $w \in \mathbb{D}$, these solutions belong to $\mathbb{D}$.
(b) If $w \in \mathbb{D}_e$, these solutions belong to $\mathbb{D}_e$.
(c) If $w \in \mathbb{T}$, these solutions belong to $\mathbb{T}$ and are distinct.

Proof By Theorem 3.3.2 we may write $B = P/P^{\#}$, where

$$P(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n, \qquad \alpha_n \neq 0, \tag{3.4.11}$$

is a polynomial of degree n whose zeros are all inside $\mathbb{D}$. Looking at the form of $P/P^{\#}$, we can divide its numerator and denominator by $|\alpha_n|$ and thus assume that $\alpha_n \in \mathbb{T}$. Let $z_1, z_2, \ldots, z_n$ denote the zeros of P, counted according to their multiplicity. Then

$$P(z) = \alpha_n(z-z_1)(z-z_2)\cdots(z-z_n).$$

Expand the preceding and compare with (3.4.11) to obtain

$$\alpha_0 = (-1)^n z_1 z_2 \cdots z_n \alpha_n,$$

which belongs to $\mathbb{D}$ since $|\alpha_n| = 1$.

The solutions to $B(z) = w$ are precisely the solutions to (3.4.11). The coefficient of z^n in (3.4.2) is nonzero if and only if $\alpha_n \neq w\overline{\alpha_0}$. If $w \in \mathbb{D}$, then $|w\overline{\alpha_0}| < |\alpha_n| = 1$ and hence $B(z) = w$ has exactly n solutions, repeated according to multiplicity. If $w \in \mathbb{D}_e$, then apply (3.1.6) and the preceding result to conclude that $B(z) = w$ has exactly n solutions in $\mathbb{D}_e$, counted according to multiplicity.

If $w \in \mathbb{T}$, then (3.1.3), (3.1.4), and (3.1.5) imply that the n solutions to $B(z) = w$ belong to $\mathbb{T}$. These solutions are distinct since Lemma 3.4.3 guarantees that B' does not vanish on $\mathbb{T}$. □

Theorem 3.4.10 says that $B(z) = w$ does not have solutions with multiplicity greater than one if $w \in \mathbb{T}$. This does not hold if $w \notin \mathbb{T}$. For example, $B(z) = z^n$ assumes the value 0 at $z = 0$ (and the value ∞ at $z = \infty$) with multiplicity n.

3.5 Unimodular Elements of the Disk Algebra

Definition 3.5.1 The *disk algebra* $\mathscr{A}(\mathbb{D})$ is the set of analytic functions on $\mathbb{D}$ that extend continuously to $\mathbb{D}^-$.

Each finite Blaschke product belongs to $\mathscr{A}(\mathbb{D})$. More generally, any rational functions with no poles in $\mathbb{D}^-$ belongs to $\mathscr{A}(\mathbb{D})$. Another example is the function defined by $\sum_{n=1}^{\infty} z^n/n^2$.

Among the elements of $\mathscr{A}(\mathbb{D})$, the finite Blaschke products can be characterized as those functions that map $\mathbb{T}$ into $\mathbb{T}$.

Theorem 3.5.2 (Fatou [46]) *If f is analytic on $\mathbb{D}$ and*

$$\lim_{|z| \to 1^-} |f(z)| = 1,$$

then f is a finite Blaschke product.

Proof Since $|f(z)| \to 1$ as $|z| \to 1^-$, there is an $r \in [0, 1)$ so that f does not vanish on the annulus $\{z : r \leqslant |z| < 1\}$. The identity principle from complex analysis asserts that f has at most a finite number of zeros in $\mathbb{D}$. Let B be a finite Blaschke product whose zeros (located in $\{z : |z| < r\}$) are precisely the zeros of f, repeated according to multiplicity. Then f/B and B/f are analytic in $\mathbb{D}$ and

$$\lim_{|z| \to 1^-} \left| \frac{f(z)}{B(z)} \right| = \lim_{|z| \to 1^-} \left| \frac{B(z)}{f(z)} \right| = 1.$$

The Maximum Modulus Principle implies that $|f/B| \leqslant 1$ and $|B/f| \leqslant 1$ on $\mathbb{D}$. Thus, f/B is constant on $\mathbb{D}$. This constant must be unimodular, so f is a unimodular scalar multiple of B. That is, f is a finite Blaschke product. □

Corollary 3.5.3 *If $f \in \mathscr{A}(\mathbb{D})$ and $f(\mathbb{T}) \subseteq \mathbb{T}$, then f is a finite Blaschke product.*

Proof If $f \in \mathscr{A}(\mathbb{D})$, then $|f|$ is continuous on $\mathbb{D}^-$. Since $\mathbb{D}^-$ is compact, $|f|$ is uniformly continuous on $\mathbb{D}^-$, so $|f(z)| \to 1$ uniformly as $|z| \to 1^-$. Now apply Theorem 3.5.2. □

There is also meromorphic version of the preceding corollary.

Corollary 3.5.4 *Suppose f is meromorphic on $\mathbb{D}$ and extends continuously to $\mathbb{T}$. If $f(\mathbb{T}) \subseteq \mathbb{T}$, then f is a quotient of two finite Blaschke products.*

Proof By hypothesis, f has finitely many poles in $\mathbb{D}$. Let B_2 be a finite Blaschke product whose zeros are precisely the poles of f in $\mathbb{D}$, repeated according to multiplicity. Then $B_1 = fB_2$ is analytic on $\mathbb{D}$ and extends continuously to $\mathbb{D}$ with $B_1(\mathbb{T}) \subseteq \mathbb{T}$. Corollary 3.5.3 implies that B_1 is a finite Blaschke product and so $f = B_1/B_2$ as required. □

3.6 Composition of Finite Blaschke Products

The family of finite Blaschke products is conformally invariant; that is, it is invariant under any change of variables $z \mapsto \varphi(z)$ for $\varphi \in \operatorname{Aut}(\mathbb{D})$. In fact, the degree of a finite Blaschke product is a conformal invariant. Recall the automorphism τ_w from (1.2.2).

Lemma 3.6.1 *Let B be a finite Blaschke product of degree n and let $w \in \mathbb{D}$. Then $\tau_w \circ B$ and $B \circ \tau_w$ are finite Blaschke products of degree n.*

Proof The function $\tau_w \circ B$ is analytic on $\mathbb{D}$, continuous on $\mathbb{D}^-$, and unimodular on $\mathbb{T}$. Corollary 3.5.3 ensures that $\tau_w \circ B$ is a finite Blaschke product. Moreover, $(\tau_w \circ B)(z) = 0$ if and only if $B(z) = w$, so Theorem 3.4.10 tells us that the equation $B(z) = w$ has exactly n solutions in $\mathbb{D}$. Thus, $\tau_w \circ B$ is a finite Blaschke product of degree n. Corollary 3.5.3 implies that $B \circ \tau_w$ is a finite Blaschke product. That its degree is n can be verified directly. □

Clearly the lemma above also holds for general $\varphi \in \operatorname{Aut}(\mathbb{D})$.

The family of all finite Blaschke products is closed under pointwise multiplication. Indeed, if we multiply two finite Blaschke products of degree n_1 and n_2, the result is a finite Blaschke product of degree $n_1 n_2$. Less obvious is that the set of finite Blaschke products is closed under composition. In fact, Lemma 3.6.1 already reveals a special case of this property.

Theorem 3.6.2 *If B_1 and B_2 are finite Blaschke products, then $B_1 \circ B_2$ is a finite Blaschke product. Moreover, if n_1 and n_2 are the degrees of B_1 and B_2, respectively, then the degree of $B_1 \circ B_2$ is $n_1 n_2$.*

Proof Denote the zeros of B_1 by $z_1, z_2, \ldots, z_n$ and write

$$B_1 = \gamma(\tau_{z_1}\tau_{z_2}\cdots\tau_{z_n}),$$

where γ is a unimodular constant. Then

$$B_1 \circ B_2 = \gamma(\tau_{z_1} \circ B_2)(\tau_{z_2} \circ B_2)\cdots(\tau_{z_n} \circ B_2).$$

By Lemma 3.6.1, each $\tau_{z_k} \circ B_2$ is a finite Blaschke product of degree n_2. Consequently, $B_1 \circ B_2$ is a finite Blaschke product of degree $n_1 n_2$. $\square$

In Chap. 9 we explore the more difficult question of when we can write a finite Blaschke product B as a composition $B = C \circ D$, in which C and D are finite Blaschke products, in a nontrivial way.

3.7 Constant Valence

Definition 3.7.1 For an analytic function $f : \mathbb{D} \to \mathbb{C}$ and $w \in \mathbb{C}$, the *valence* of f at w is

$$v_f(w) = |\{z \in \mathbb{D} : f(z) = w\}|,$$

where $|E|$ denotes the cardinality of a set E.

Notice that $v_f(w) \in \mathbb{N} \cup \{0, \infty\}$. Theorem 3.4.10 says that for a finite Blaschke product B of degree n,

$$v_B(w) = n, \quad w \in \mathbb{D}.$$

This constant valence property characterizes the finite Blaschke products of degree n amongst the functions in the Schur class $\mathscr{S}$ [43–45].

Theorem 3.7.2 (Fatou) *Let $f \in \mathscr{S}$ and $n \in \mathbb{N}$ such that with $v_f(w) = n$ for all $w \in \mathbb{D}$. Then f is a finite Blaschke product of degree n.*

Proof We follow the proof of Radó [116] and show that

$$\lim_{|z| \to 1^-} |f(z)| = 1. \tag{3.7.3}$$

If we can do this, then Theorem 3.5.2 would imply that f is a finite Blaschke product. Suppose toward a contradiction that $f : \mathbb{D} \to \mathbb{D}$ is an analytic function of constant valence $n \geqslant 1$ and that (3.7.3) fails. Then there is a sequence z_m of distinct points in $\mathbb{D}$ and a $w_0 \in \mathbb{D}$ so that

$$\lim_{m \to \infty} |z_m| = 1 \quad \text{and} \quad \lim_{m \to \infty} f(z_m) = w_0. \tag{3.7.4}$$

Indeed, let z_m be a sequence of distinct points in $\mathbb{D}$ for which $|z_m| \to 1$ and $|f(z_m)|$ remains bounded away from 1. Passing to a subsequence and relabeling, we may assume that there is an $r \in (0,1)$ so that $|f(z_m)| \leqslant r < 1$ for all m. The compactness of $|z| \leqslant r$ implies that a further subsequence, which we relabel as z_m for convenience, satisfies (3.7.4) for some $w_0 \in \mathbb{D}$.

Since f has constant valence n, it follows that $f(z_m) \neq w_0$ for all but finitely many m. Let $a_1, a_2, \ldots, a_k$ be the distinct solutions of $f(z) = w_0$, respectively, with multiplicities $n_1, n_2, \ldots, n_k$. By assumption,

$$n_1 + n_2 + \cdots + n_k = n.$$

About each point a_j, the function f has a power series expansion

$$f(z) = w_0 + \sum_{k=n_j}^{\infty} \frac{f^{(k)}(a_j)}{k!}(z - a_j)^k,$$

in which $f^{(n_j)}(a_j) \neq 0$. If $\epsilon_j > 0$ is sufficiently small, we can write

$$f(z) = w_0 + \big((z - a_j) f_j(z)\big)^{n_j} \tag{3.7.5}$$

for z contained in

$$D(a_j, \epsilon_j) = \{z : |z - a_j| < \epsilon_j\},$$

in which f_j is a nonvanishing analytic function on $D(a_j, \epsilon_j)$. Without loss of generality, we impose the extra restrictions

$$\epsilon_j < \min\left\{\tfrac{1}{2}|a_j - a_i| : 1 \leqslant i \leqslant k,\ i \neq j\right\} \quad \text{and} \quad \epsilon_j < \tfrac{1}{2}(1 - |a_j|)$$

to ensure that the disks $D(a_j, \epsilon_j)$ are pairwise disjoint and do not intersect $\mathbb{T}$.

Since $g_j(z) = (z - a_j) f_j(z)$ has a simple zero at a_j, it is injective on a small neighborhood of a_j. Thus, if necessary, we can make each ϵ_j even smaller so that $g_j(z)$ is injective on $D(a_j, \epsilon_j)$. The Open Mapping Theorem says that

$$\bigcap_{j=1}^{k} g_j(D(a_j, \epsilon_j))$$

is an open set that contains the origin. Let $\epsilon > 0$ be small enough so that

$$D(0, \epsilon) \subseteq \bigcap_{j=1}^{k} g_j\big(D(a_j, \epsilon_j)\big)$$

and let

$$V_j = g_j^{-1}\big(D(0, \epsilon)\big) \subseteq D(a_j, \epsilon_j), \qquad \text{for } 1 \leqslant j \leqslant k.$$

Observe that $g_j : V_j \to D(0, \epsilon)$ is bijective, $g_j(a_j) = 0$, and the open sets V_j are pairwise disjoint and do not intersect $\mathbb{T}$. Consequently, (3.7.5) tells us that for each $w \in D(w_0, \epsilon')\backslash\{w_0\}$, where

$$\epsilon' = \epsilon^{\max\{n_1,\ldots,n_k\}},$$

the equation $f(z) = w$ has exactly n_j distinct solutions in V_j for each j.

Since $f(z_m) \to w_0$ and $|z_m| \to 1$, for sufficiently large m we have

$$f(z_m) \in D(w_0, \epsilon'), \qquad f(z_m) \neq w_0, \quad \text{and} \quad z_m \notin \bigcup_{j=1}^{k} V_j.$$

Fix any such m, and let $w_m = f(z_m)$. Then each V_j contains n_j distinct points that map to w_m. Thus, the equation $f(z) = w_m$ has at least

$$n_1 + n_2 + \cdots + n_k + 1 = n + 1$$

solutions. This is a contradiction. □

3.8 Finite Blaschke Products on $\mathbb{C}_+$

Recall that

$$\mathbb{C}_+ = \{z \in \mathbb{C} : \operatorname{Im} z > 0\} \quad \text{and} \quad \mathbb{C}_- = \{z \in \mathbb{C} : \operatorname{Im} < 0\}$$

denote the upper and lower half planes, respectively. If $z_1, z_2, \ldots, z_n \in \mathbb{C}_+$ and $\gamma \in \mathbb{T}$, then

$$B(z) = \gamma \prod_{k=1}^{n} \frac{z - z_k}{z - \overline{z_k}} \tag{3.8.1}$$

is a finite *Blaschke product of degree n for* $\mathbb{C}_+$. If $\varphi : \mathbb{D} \to \mathbb{C}_+$ is the Möbius transformation

$$\varphi(z) = i\frac{1+z}{1-z},$$

(see (2.5.1) and Fig. 2.4), then B is a finite Blaschke product for $\mathbb{C}_+$ if and only if $B \circ \varphi$ is a finite Blaschke product for $\mathbb{D}$. In light of this relationship, we mostly consider finite Blaschke products on $\mathbb{D}$.

The following properties of (nonconstant) finite Blaschke products on the $\mathbb{C}_+$ follow directly from the corresponding properties (3.1.3), (3.1.4), (3.1.5), and (3.1.6) of finite Blaschke products on $\mathbb{D}$:

$$|B(z)| < 1, \quad \text{if } z \in \mathbb{C}_+, \tag{3.8.2}$$

$$|B(x)| = 1, \quad \text{if } x \in \mathbb{R}, \tag{3.8.3}$$

$$|B(z)| > 1, \quad \text{if } z \in \mathbb{C}_-, \tag{3.8.4}$$

and

$$B(z) = \frac{1}{\overline{B(\overline{z})}}, \quad z \in \widehat{\mathbb{C}}. \tag{3.8.5}$$

An important distinction between finite Blaschke products on $\mathbb{D}$ and $\mathbb{C}_+$ concerns the locations of their poles. A nonconstant finite Blaschke product for the upper half plane has at least one finite pole; this is evident in the definition (3.8.1). In contrast, the nonconstant finite Blaschke products $z, z^2, z^3, \ldots$ on $\mathbb{D}$ are entire functions. This difference is important to remember in certain applications.

3.9 Notes

Commuting Blaschke Products

From Theorem 3.6.2 we know that if B_1 and B_2 are finite Blaschke products, then $B_1 \circ B_2$ is another finite Blaschke product. However, see Exercise 3.6, it is not always the case that $B_1 \circ B_2 = B_2 \circ B_1$. The paper [19] explores when B_1 and B_2 commute.

Infinite Blaschke Products

One can extend the notion of finite Blaschke products to *infinite* Blaschle products, where the number of factors is infinite. As expected, there are convergence issues. Indeed, for an infinite sequence z_n of points in $\mathbb{D}\backslash\{0\}$, we define the formal product

$$B(z) = \prod_{n=1}^{\infty} \frac{|z_n|}{z_n} \frac{z_n - z}{1 - \overline{z_n} z}. \tag{3.9.1}$$

The product above converges uniformly on compact subsets of $\mathbb{D}$ if and only if the zeros z_n satisfy the *Blaschke condition*

$$\sum_{n=1}^{\infty} (1 - |z_n|) < \infty.$$

If this occurs, then B defines a bounded analytic function on $\mathbb{D}$ such that $|B| \leqslant 1$ on $\mathbb{D}$ and such that

$$\lim_{r \to 1^-} B(re^{i\theta})$$

exists for almost every θ. Furthermore, this value is unimodular almost everywhere [25, 38].

There are beautiful theorems of Frostman [50] that discuss the behavior of B on $\mathbb{T}$. For a fixed θ, the radial limits of a Blaschke product (and all of its subproducts) exist and have modulus equal to one at $e^{i\theta}$ if and only if

$$\sum_{n=1}^{\infty} \frac{1 - |a_n|}{|e^{i\theta} - a_n|} < \infty.$$

Furthermore, the radial limit of B' exists at $e^{i\theta}$ if and only if

$$\sum_{n=1}^{\infty} \frac{1 - |a_n|^2}{|e^{i\theta} - a_n|^2} < \infty$$

and, moreover, $|B'(e^{i\theta})|$ is equal to the sum above. Compare this to (3.4.8).

By the argument used to prove Corollary 8.2.6 below, one can show that if

$$B(z) = \sum_{n=0}^{\infty} c_n z^n,$$

then $c_n = o(1)$. Furthermore, an argument used to prove Corollary 10.1.16 below shows that for a finite Blaschke product we have $c_n = o(1/n)$. A theorem of Shapiro and Newman [107] shows that a general Blaschke product is a finite Blaschke product if and only if $c_n = o(1/n)$. Thus, in terms of Taylor coefficients, we have a characterization of the finite Blaschke products among the set of all Blaschke products.

Finally, we mention that Blaschke products (finite or infinite) are special cases of a wider class of analytic functions on $\mathbb{D}$ called *inner functions* [38, 61].

3.10 Exercises

3.1 Prove that if f is analytic on $\mathbb{D}$, then $g(z) = \overline{f(\overline{z})}$ is also analytic on $\mathbb{D}$.

3.2 Prove that if B_1 and B_2 are finite Blaschke products whose degrees do not exceed n and $B_1 = B_2$ at $n + 1$ points of $\mathbb{D}$, then $B_1 \equiv B_2$.
Hint: If $B_1 = B_2$ at $n + 1$ points, then, considering B_1 and B_2 as meromorphic functions on $\mathbb{C}$, they are equal at $2n + 2$ points (which ones?).

3.3 In [108] the authors compare results about polynomials with some results about finite Blaschke products. For some examples of this, prove the following.

(a) Let f be entire with $f(\mathbb{C}) = \mathbb{C}$. Then f is a polynomial of degree n if and only if the valence of f is n at each point of $\mathbb{C}$. This is the polynomial analogue of Theorem 3.7.2.

(b) If f is entire and

$$\lim_{|z|\to\infty} |f(z)| = \infty,$$

then f is a polynomial. This is the analogue of Theorem 3.5.2.

(c) If p and q are polynomials whose degrees do not exceed n and if $p = q$ at $n+1$ distinct points, then $p \equiv q$. This is the analogue of Exercise 3.2.

(d) For a given $w \in \mathbb{C}$, p is a polynomial of degree n if and only if

$$\frac{p(z) - p(w)}{z - w}$$

is also a polynomial of degree $n-1$. In [5], they prove the (hyperbolic) analogue of this for finite Blaschke products. Indeed, for $z, w \in \mathbb{D}$, define

$$[z, w] = \frac{z - w}{1 - \overline{w}z}.$$

One can show that if $w \in \mathbb{D}$, then B is a Blaschke product of degree n if and only if

$$\frac{[B(z), B(w)]}{[z, w]}$$

is a Blaschke product of degree $n - 1$.

3.4 Prove (3.1.6) by direct computation.

3.5 Let B be a finite Blaschke product of order n and α_1, α_2 be distinct points on $\mathbb{T}$. If $\zeta_1, \zeta_2, \ldots, \zeta_n$ are the solutions to $B(z) = \alpha_1$ and $\xi_1, \xi_2, \ldots, \xi_n$ are the solutions to $B(z) = \alpha_2$, show that the ζ_js alternate with the ξ_js as one travels around $\mathbb{T}$.

3.6 Produce an example of two finite Blaschke products B_1, B_2 for which $B_1 \circ B_2 \neq B_2 \circ B_1$. See the end notes of this chapter for more on this.

3.7 (a) Use Corollary 3.5.3 to provide another proof of Theorem 3.3.2. (b) Use Corollary 3.5.4 to provide another proof of Corollary 3.3.10.

3.8 Let $z_0 \in \mathbb{D}_+$, where $\mathbb{D}_+ = \mathbb{D} \cap \mathbb{C}_+$, and set

$$f(z) = \frac{(z_0 - z)\,(1 - z_0 z)}{(\overline{z}_0 - z)\,(1 - \overline{z}_0 z)}, \qquad z \in \mathbb{D}_+.$$

Show that

$$|f(z)| < 1, \qquad z \in \mathbb{D}_+,$$

and

$$|f(z)| = 1, \qquad z \in \partial\mathbb{D}_+.$$

Hint: We have

$$f = \frac{\overline{z}_0}{z_0} \cdot \frac{b_{z_0}}{b_{\overline{z}_0}}.$$

3.9 Let $z_0 \in \mathbb{D}_+$ and let

$$D_{z_0} = \left\{z : |z| < 1,\ \operatorname{Im} z \geqslant 0,\ |z-1| \leqslant |z_0 - 1|\right\}.$$

Show that

$$\left|\frac{1 - z_0\, z}{\overline{z}_0 - z}\right| \leqslant \exp\left(\frac{2}{\sin^2 \vartheta_0}\right),$$

in which $\vartheta_0 = -\arg(1 - z_0) \in (0, \frac{\pi}{2})$.
Hint: Use the fact that $\log x \leqslant x - 1$ for $x \geqslant 1$ and apply it to

$$\log\left|\frac{1 - z_0\, z}{\overline{z}_0 - z}\right|^2.$$

The identity (1.6.1) may be helpful.

3.10 Let $r \in [0, 1)$ and $z \in \mathbb{D}^-$.

(a) Prove that $\left|\dfrac{1-z}{1-rz}\right| \leqslant \dfrac{2}{1+r}$.

(b) Prove that $\dfrac{1}{|1-rz|} \leqslant \dfrac{2}{|1-z|}$.

3.11 Show that

$$\frac{1}{|1 - re^{-i\theta}\, z|} \leqslant \frac{2}{|e^{i\theta} - z|}$$

for all $z,\ re^{i\theta} \in \mathbb{D}$.
Hint: Use Exercise 3.10.

3.12 Let f be an entire function of constant modulus on the unit circle $\mathbb{T}$. Show that $f(z) = cz^n$, in which c is a constant and n is a nonnegative integer.
Hint: Apply Corollary 3.5.3.

3.13 Let Ω_1 and Ω_2 be two bounded regions in $\mathbb{C}$ and let $f : \Omega_1 \to \Omega_2$ be analytic. Suppose that there is no sequence z_n in Ω_1 that converges to a point on $\partial\Omega_1$ with $f(z_n)$ in Ω_2 converges to a point on Ω_2. Show that f has a constant valence on Ω_1.

Remark The main assumption of this result is equivalent to each of the following.

(a) For each compact set $K \subseteq \Omega_2$, the set $f^{-1}(K)$ is a compact subset of Ω_1.
(b) If $E \subseteq \Omega_2$ is such that $\operatorname{dist}(E, \partial\Omega_2) > 0$, then $\operatorname{dist}(f^{-1}(E), \partial\Omega_1) > 0$.

3.14 (Carathéodory–Rademacher [17]) Let Ω_1 and Ω_2 be two bounded regions in $\mathbb{C}$, and let $f : \Omega_1 \to \Omega_2$ be analytic. Suppose that f has constant valence on Ω_2. Show that there is no sequence z_n in Ω_1 that converges to a point in $\partial\Omega_1$ and at the same time the sequence $f(z_n)$ in Ω_2 converges to a point inside Ω_2.
Hint: The proof of Theorem 3.7.2 may help.

Chapter 4
Approximation by Finite Blaschke Products

Although finite Blaschke products are a remarkable and exclusive class of functions, they appear in many important approximation problems.

Let H^∞ denote the set of all bounded analytic functions on $\mathbb{D}$. Since H^∞ is closed under addition and scalar multiplication, it is a vector space. It is also closed under pointwise multiplication, so H^∞ is an algebra over $\mathbb{C}$. We endow H^∞ with the norm

$$\|f\|_\infty := \sup\{|f(z)| : z \in \mathbb{D}\}. \tag{4.0.1}$$

With respect to this norm, H^∞ is a Banach algebra: it is a normed algebra that is complete with respect to the metric induced by the norm (4.0.1), which is submultiplicative:

$$\|fg\|_\infty \leqslant \|f\|_\infty \|g\|_\infty.$$

The closed unit ball $\{f \in H^\infty : \|f\|_\infty \leqslant 1\}$ of H^∞ is the Schur class $\mathscr{S}$ (Definition 1.0.2).

4.1 Approximating Functions from $\mathscr{S}$

If an analytic function f on $\mathbb{D}$ can be uniformly approximated on $\mathbb{D}$ by a sequence of finite Blaschke products, then it is uniformly continuous on $\mathbb{D}$. Consequently, f has a unique continuous extension to $\mathbb{D}^-$ and, moreover, this extension is unimodular on $\mathbb{T}$. Thus, f is itself a finite Blaschke product (Corollary 3.5.3). In other words, the set of finite Blaschke products is a proper, norm-closed subset of $\mathscr{S}$ with respect to the norm (4.0.1).

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_4

A generic element of $\mathscr{S}$ need not have a continuous extension to $\mathbb{D}^-$. For example, consider the function

$$f(z) = \exp\Big(-\frac{1+z}{1-z}\Big);$$

see Exercise 4.1. Consequently, we cannot expect to approximate this function by a sequence of finite Blaschke products in the norm (4.0.1). On the other hand, if we consider the topology of uniform convergence on compact subsets of $\mathbb{D}$, then the finite Blaschke products are dense in $\mathscr{S}$.

Theorem 4.1.1 (Carathéodory [16]) *For each $f \in \mathscr{S}$, there is a sequence of finite Blaschke products that converges uniformly on compact subsets of $\mathbb{D}$ to f.*

Proof It suffices to show that for each $f \in \mathscr{S}$ and $n \geqslant 1$, there is a finite Blaschke product B_n so that $f - B_n$ has a zero of order at least n at the origin. If this occurs, then $f - B_n = z^{n-1}g$, where $g \in H^\infty$, $g(0) = 0$, and

$$\|g\|_\infty = \|z^{n-1}g\|_\infty = \|f - B_n\|_\infty \leqslant 2.$$

Thus, by the Schwarz Lemma (Lemma 1.1.1),

$$|f(z) - B_n(z)| \leqslant 2|z|^n, \quad z \in \mathbb{D}.$$

The preceding inequality shows that $B_n \to f$ uniformly on compact subsets of $\mathbb{D}$.

To show that for each $f \in \mathscr{S}$ and $n \geqslant 1$, there is a finite Blaschke product B_n so that $f - B_n$ has a zero of order at least n at the origin, we proceed by induction on n. Our base case is $n = 1$. For each $f \in \mathscr{S}$, we have $c_0 = f(0) \in \mathbb{D}^-$. If $|c_0| = 1$, then the Maximum Modulus Principle says that f is a unimodular constant and there is nothing to prove. If $|c_0| < 1$, then $B_0(z) = \tau_{c_0}(z)$ is a finite Blaschke product with the same constant term as f and thus $f - B_0$ vanishes at the origin.

Suppose for our induction hypothesis that for each $f \in \mathscr{S}$ there is a finite Blaschke product B_n so that $f - B_n$ has a zero of order at least n at the origin. Since $f \in \mathscr{S}$, the Schwarz Lemma implies that

$$g(z) = \frac{\tau_{c_0}(f(z))}{z}$$

belongs to $\mathscr{S}$. Hence there is a finite Blaschke product B_n so that $g - B_n$ has a zero of order at least n at the origin. In other words,

$$g(z) - B_n(z) = z^n K(z), \quad z \in \mathbb{D}, \tag{4.1.2}$$

for some $K \in H^\infty$. Since

$$f(z) = \tau_{c_0}(zg(z)),$$

and since

$$B_{n+1}(z) = \tau_{c_0}(zB_n(z))$$

is a finite Blaschke product by Lemma 3.6.1, we expect that B_{n+1} has the desired properties. To establish this, first observe that

$$\tau_{c_0}(z_2) - \tau_{c_0}(z_1) = \frac{(1-|c_0|^2)(z_1 - z_2)}{(1-\overline{c}_0 z_1)(1-\overline{c}_0 z_2)} \tag{4.1.3}$$

for $z_1, z_2 \in \mathbb{D}$; see Exercise 4.3. Then conclude that

$$\begin{aligned} f(z) - B_{n+1}(z) &= \tau_{c_0}\big(zg(z)\big) - \tau_{c_0}\big(zB_n(z)\big) \\ &= \frac{(1-|c_0|^2)(zg(z) - zB_n(z))}{(1-\overline{c}_0 zg(z))(1-\overline{c}_0 zB_n(z))} && \text{(by (4.1.3))} \\ &= H(z)z(g(z) - B_n(z)) \\ &= z^{n+1}K(z)H(z), && \text{(by (4.1.2))} \end{aligned}$$

where

$$H(z) = \frac{1-|c_0|^2}{(1-\overline{c_0}zg(z))(1-\overline{c_0}zB_n(z))}$$

and K is the function from (4.1.2). This completes the induction and the proof. □

4.2 The Closed Convex Hull of the Finite Blaschke Products

Definition 4.2.1 Let $\mathscr{V}$ be a vector space. A *convex combination* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathscr{V}$ is an expression of the form

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n,$$

in which

$$\lambda_1, \lambda_2, \ldots, \lambda_n \in [0, 1] \quad \text{and} \quad \lambda_1 + \lambda_2 + \cdots + \lambda_n = 1.$$

The set

$$\operatorname{conv}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$$

of all convex combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is the *convex hull* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

Since each finite Blaschke product belongs to

$$\{f \in \mathscr{A}(\mathbb{D}) : \|f\|_\infty \leqslant 1\},$$

the closed unit ball of the disk algebra $\mathscr{A}(\mathbb{D})$ (recall Definition 3.5.1). Any convex combination of finite Blaschke products belongs to the unit ball of $\mathscr{A}(\mathbb{D})$. In fact, they are dense in $\mathscr{A}(\mathbb{D})$.

Theorem 4.2.2 (Fisher [48]) *Each function in the closed unit ball of $\mathscr{A}(\mathbb{D})$ can be uniformly approximated on $\mathbb{D}^-$ by a sequence of convex combinations of finite Blaschke products.*

Proof Fix f in the closed unit ball of $\mathscr{A}(\mathbb{D})$ and $\epsilon > 0$. Since f is continuous on $\mathbb{D}^-$, uniform continuity implies that there is a $t \in [0, 1)$ such that

$$\|f_t - f\|_\infty < \frac{\epsilon}{2},$$

where $f_t(z) = f(tz)$ is a dilation of f. Theorem 4.1.1 provides a finite Blaschke product B so that

$$\sup_{z \in t\mathbb{D}} |f(z) - B(z)| < \frac{\epsilon}{2}.$$

Since

$$\begin{aligned} \|f_t - B_t\|_\infty &= \sup_{z \in \mathbb{D}} |f_t(z) - B_t(z)| \\ &= \sup_{z \in t\mathbb{D}} |f(z) - B(z)| \\ &< \frac{\epsilon}{2}, \end{aligned}$$

the finite Blaschke product B satisfies

$$\begin{aligned} \|f - B_t\|_\infty &\leqslant \|f - f_t\|_\infty + \|f_t - B_t\|_\infty \\ &\leqslant \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

If we can show that B_t is itself a convex combination of finite Blaschke products, the proof will be complete.

To accomplish this, first observe that the product of two convex combinations of finite Blaschke products is a convex combination of finite Blaschke products; see Exercise 4.5. Since $(gh)_t = g_t h_t$ for any analytic functions g and h on $\mathbb{D}$, we may assume that the finite Blaschke product B takes the form

$$B(z) = \frac{\alpha - z}{1 - \overline{\alpha} z}, \quad \alpha \in \mathbb{D}.$$

A computation confirms that

$$B_t(z) = \frac{t(1-|\alpha|^2)}{1-|\alpha|^2t^2} \cdot \frac{\alpha t - z}{1-\overline{\alpha}tz} + \frac{|\alpha|(1-t^2)}{1-|\alpha|^2t^2} \cdot e^{i\arg\alpha}; \tag{4.2.3}$$

see Exercise 4.6. The expression on the right-hand side of (4.2.3) is almost what we want; it is a linear combination, with positive coefficients, of a disk automorphism and a unimodular constant (which is a finite Blaschke product). However, the coefficients need not sum to one. Fortunately,

$$1 - \frac{t(1-|\alpha|^2)}{1-|\alpha|^2t^2} - \frac{|\alpha|(1-t^2)}{1-|\alpha|^2t^2} = \frac{(1-t)(1-|\alpha|)}{1+|\alpha|t} > 0, \tag{4.2.4}$$

so we add

$$0 = \frac{(1-t)(1-|\alpha|)}{2(1+|\alpha|t)} \cdot 1 + \frac{(1-t)(1-|\alpha|)}{2(1+|\alpha|t)} \cdot (-1)$$

to both sides of (4.2.3) and obtain

$$\begin{aligned} B_t(z) = {} & \frac{t(1-|\alpha|^2)}{1-|\alpha|^2t^2} \cdot \frac{\alpha t - z}{1-\overline{\alpha}tz} + \frac{|\alpha|(1-t^2)}{1-|\alpha|^2t^2} \cdot e^{i\arg\alpha} \\ & + \frac{(1-t)(1-|\alpha|)}{2(1+|\alpha|t)} \cdot 1 + \frac{(1-t)(1-|\alpha|)}{2(1+|\alpha|t)} \cdot (-1), \end{aligned}$$

which is a convex combination of four finite Blaschke products (three of which are unimodular constants). □

4.3 Approximating Continuous Unimodular Functions

If B_1 and B_2 are finite Blaschke products, then the restriction of B_1/B_2 to $\mathbb{T}$ is a continuous unimodular function. If u is the boundary function, with unimodular values, for a meromorphic function on $\mathbb{D}$ with a finite number of zeros and poles in $\mathbb{D}$, then Corollary 3.5.4 says that u can be approximated uniformly on $\mathbb{T}$ by unimodular functions of the form B_1/B_2, in which B_1, B_2 are finite Blaschke products. The main result of this section is an improvement of this fact.

Theorem 4.3.1 (Helson–Sarason [77]) *Let $u : \mathbb{T} \to \mathbb{T}$ be continuous and let $\epsilon > 0$. Then there are finite Blaschke products B_1 and B_2 such that*

$$\max_{\xi\in\mathbb{T}} \left| u(\xi) - \frac{B_1(\xi)}{B_2(\xi)} \right| < \epsilon.$$

The proof of this theorem requires the following technical lemma which is a precise formulation of the fact that a closed curve that does not pass through the origin winds around the origin either an even number or an odd number of times.

Lemma 4.3.2 *Let $\gamma : \mathbb{T} \to \mathbb{T}$ be continuous. Then there exists a continuous unimodular function η on $\mathbb{T}$ such that either*

$$\gamma(\zeta) = \eta^2(\zeta), \qquad \zeta \in \mathbb{T},$$

or

$$\gamma(\zeta) = \zeta\eta^2(\zeta), \qquad \zeta \in \mathbb{T}.$$

Proof Since $\gamma : \mathbb{T} \to \mathbb{T}$ is uniformly continuous, there is a positive integer N so that

$$|\theta - \theta'| \leqslant \frac{2\pi}{N} \qquad \Longrightarrow \qquad |\gamma(e^{i\theta}) - \gamma(e^{i\theta'})| < 2. \tag{4.3.3}$$

Divide $\mathbb{T}$ into N arcs

$$T_k = \left\{ e^{i\theta} : \frac{2(k-1)\pi}{N} \leqslant \theta \leqslant \frac{2k\pi}{N} \right\}, \qquad 1 \leqslant k \leqslant N,$$

of equal length. The condition (4.3.3), along with the continuity of γ and the compactness of T_k, imply that $\gamma(T_k)$ is a closed sub-arc of $\mathbb{T}$ that subtends an angle strictly less than π. Thus, there is a continuous function $\phi_k(\theta)$, defined for

$$\theta \in \left[\frac{2(k-1)\pi}{N}, \frac{2k\pi}{N}\right],$$

such that

$$\gamma(e^{i\theta}) = \exp(i\phi_k(\theta)), \qquad e^{i\theta} \in T_k.$$

The ϕ_k are uniquely defined up to an additive multiple of 2π. Since γ is continuous, we adjust those additive constants so that

$$\phi_k\left(\frac{2k\pi}{N}\right) = \phi_{k+1}\left(\frac{2k\pi}{N}\right), \qquad 1 \leqslant k \leqslant N-1. \tag{4.3.4}$$

Define $\phi : [0, 2\pi] \to \mathbb{R}$ by

$$\phi(\theta) = \phi_k(\theta), \quad \theta \in \left[\frac{2(k-1)\pi}{N}, \frac{2k\pi}{N}\right], \quad k = 1, 2, \ldots, N.$$

By (4.3.4), we obtain a continuous function on $[0, 2\pi]$ such that

$$\gamma(e^{i\theta}) = \exp(i\phi(\theta)), \qquad e^{i\theta} \in \mathbb{T}.$$

The continuity of γ at 1 implies that $\phi(2\pi) - \phi(0)$ is an integer multiple of 2π. If

$$\frac{\phi(2\pi) - \phi(0)}{2\pi}$$

is even, let

$$\eta(e^{i\theta}) = \exp\left(\frac{i\phi(\theta)}{2}\right);$$

if it is odd, let

$$\eta(e^{i\theta}) = \exp\left(i\frac{1}{2}(\phi(\theta) - \theta)\right).$$

Then η is continuous and unimodular on $\mathbb{T}$ and it satisfies either $\gamma(e^{i\theta}) = \eta^2(e^{i\theta})$ for all θ or $\gamma(e^{i\theta}) = e^{i\theta}\,\eta^2(e^{i\theta})$ for all θ. □

Proof (of Theorem 4.3.1) By Lemma 4.3.2, it suffices to prove our claim for unimodular functions of the form $\gamma = \eta^2$. This is because $e^{i\theta}$ is the boundary function for the finite Blaschke product z. Without loss of generality, we may assume that $0 < \epsilon < 1$.

By the Stone–Weierstrass Theorem, there is a trigonometric polynomial

$$h(e^{i\theta}) = \sum_{n=-N}^{N} a_n e^{in\theta}$$

such that

$$\max_{\xi \in \mathbb{T}} |\eta(\xi) - h(\xi)| < \frac{\epsilon}{2}.$$

Since $0 < \epsilon < 1$ and η is unimodular, the preceding inequality implies that h has no zeros on $\mathbb{T}$. Let

$$h^*(z) = \overline{h(1/\overline{z})}, \qquad z \in \mathbb{T},$$

and

$$f(z) = \frac{h(z)}{h^*(z)}.$$

For $\zeta \in \mathbb{T}$,

$$\begin{aligned} |f(\zeta)| &= \left|\frac{h(\zeta)}{h^*(\zeta)}\right| \\ &= \frac{|h(\zeta)|}{|h(1/\overline{\zeta})|} \\ &= \frac{|h(\zeta)|}{|h(\zeta)|} \\ &= 1, \end{aligned}$$

that is, f is unimodular on $\mathbb{T}$. Moreover, on $\mathbb{T}$,

$$\begin{aligned}
|\gamma - f| &= |\eta^2 - h/h^*| \\
&= \left|\frac{\eta}{\eta^*} - \frac{h}{h^*}\right| \\
&= \left|\frac{(\eta - h)h^* + (h^* - \eta^*)h}{\eta^* h^*}\right| \\
&\leqslant \frac{|\eta - h||h^*| + |h^* - \eta^*||h|}{|\eta^*||h^*|} \\
&= \frac{|\eta - h||h| + |\eta - h||h|}{|\eta||h|} \\
&= 2|\eta - h| \\
&< 2 \cdot \frac{\epsilon}{2} \cdot 1 \\
&= \epsilon.
\end{aligned}$$

Since f is a meromorphic function that is unimodular and continuous on $\mathbb{T}$, Corollary 3.5.4 implies that it is the quotient of two finite Blaschke products. This concludes the proof. □

4.4 Approximation by Finite Blaschke Products with Simple Zeros

The finite Blaschke products produced by Theorems 4.1.1, 4.2.2, and 4.3.1 might have repeated zeros. In particular applications, one might require the approximating finite Blaschke products to have simple zeros. The following theorem remedies this situation.

Theorem 4.4.1 *Let B be a finite Blaschke product of degree n. Then there is a family of finite Blaschke products $\{B_\epsilon \colon 0 < \epsilon < \epsilon_0\}$ with the following properties.*

(a) B_ϵ is of degree n.
(b) Each B_ϵ has distinct zeros.
(c) For each ϵ, $B_\epsilon(0) \neq 0$ and $B_\epsilon'(0) \neq 0$.
(d) As $\epsilon \to 0$, B_ϵ converges uniformly to B on any compact subset of $\mathbb{C}$ that does not contain a pole of B. In particular, B_ϵ converges uniformly to B on $\mathbb{D}^-$.

For a generalization of the preceding theorem, see Exercise 4.7. The proof of Theorem 4.4.1 requires the following lemma.

Lemma 4.4.2 *Let*

$$B(z) = \prod_{k=1}^{n} \frac{z_k - z}{1 - \overline{z_k} z},$$

and suppose that we have a family of finite Blaschke products

$$B_\epsilon(z) = \prod_{k=1}^{n} \frac{z_{k,\epsilon} - z}{1 - \overline{z_{k,\epsilon}} z}, \qquad 0 < \epsilon < \epsilon_0,$$

such that

$$\lim_{\epsilon \to 0} z_{k,\epsilon} = z_k, \qquad 1 \leqslant k \leqslant n. \tag{4.4.3}$$

Then B_ϵ converges uniformly to B as $\epsilon \to 0$ on all compact subsets of $\mathbb{C}$ that do not contain a pole of B. In particular, B_ϵ converges uniformly to B on $\mathbb{D}^-$.

Proof If $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are two finite sequences in $\mathbb{D}^-$, then

$$\begin{aligned} \left| \prod_{k=1}^{n} a_k - \prod_{k=1}^{n} b_k \right| &= \left| \prod_{k=1}^{n} a_k - b_1 \prod_{k=2}^{n} a_k + b_1 \prod_{k=2}^{n} a_k - \prod_{k=1}^{n} b_k \right| \\ &= \left| (a_1 - b_1) \prod_{k=2}^{n} a_k + b_1 \left(\prod_{k=2}^{n} a_k - \prod_{k=2}^{n} b_k \right) \right| \\ &\leqslant |a_1 - b_1| \prod_{k=2}^{n} |a_k| + |b_1| \left| \prod_{k=2}^{n} a_k - \prod_{k=2}^{n} b_k \right| \\ &\leqslant |a_1 - b_1| + \left| \prod_{k=2}^{n} a_k - \prod_{k=2}^{n} b_k \right| \end{aligned}$$

and hence induction yields

$$\left| \prod_{k=1}^{n} a_k - \prod_{k=1}^{n} b_k \right| \leqslant \sum_{k=1}^{n} |a_k - b_k|.$$

Therefore,

$$|B_\epsilon(z) - B(z)| \leqslant \sum_{k=1}^{n} \left| \frac{z_{k,\epsilon} - z}{1 - \overline{z_{k,\epsilon}} z} - \frac{z_k - z}{1 - \overline{z_k} z} \right|. \tag{4.4.4}$$

Fix a compact set K in $\mathbb{C}$ that does not contain any pole of B. The assumption (4.4.3) ensures that the set of poles of B_ϵ is eventually disjoint from K. For $z \in K$ and ϵ sufficiently small, $1-\overline{z_{k,\epsilon}}z$ and $1-\overline{z_k}z$ are bounded away from zero. Consequently,

$$\left| \frac{z_{k,\epsilon}-z}{1-\overline{z_{k,\epsilon}}z} - \frac{z_k-z}{1-\overline{z_k}z} \right| \leqslant M\,|z_{k,\epsilon}-z_k| \tag{4.4.5}$$

for some constant M that depends on K; see Exercise 4.8. From (4.4.4) and (4.4.5) we deduce that

$$|B_\epsilon(z)-B(z)| \leqslant M\sum_{k=1}^{n}|z_{k,\epsilon}-z_k|, \qquad z\in K.$$

Thus, B_ϵ converges uniformly to B on K. □

Proof (of Theorem 4.4.1) Write

$$B(z) = e^{i\beta}z^{j_0}\prod_{k=1}^{m}\left(\frac{z_k-z}{1-\overline{z_k}z}\right)^{j_k},$$

in which $\beta\in[0,2\pi)$, $j_k\geqslant 1$, and $z_1, z_2, \ldots, z_m$ are distinct elements of $\mathbb{D}\backslash\{0\}$. Let $z_0=0$ and define

$$\epsilon_0 = \min\left\{\tfrac{1}{2}|z_k-z_\ell| : 0\leqslant k,\ell\leqslant m,\ k\neq \ell\right\},$$

which is positive. If $0<\epsilon<\epsilon_0$, then the circles

$$\Gamma_{k,\epsilon} = \{z\in\mathbb{C} : |z-z_k| = \epsilon\}$$

are pairwise disjoint and do not pass through the origin. On each $\Gamma_{k,\epsilon}$, consider any arbitrary set of distinct j_k elements, say

$$z_{k,\epsilon,1}, \quad z_{k,\epsilon,2}, \ldots, \quad z_{k,\epsilon,j_k}, \qquad 0\leqslant k\leqslant m, \tag{4.4.6}$$

and form the finite Blaschke product

$$B_\epsilon(z) = e^{i\beta}(-1)^{j_0}\prod_{k=0}^{m}\prod_{\ell=0}^{j_k}\frac{z_{k,\epsilon,\ell}-z}{1-\overline{z_{k,\epsilon,\ell}}z}.$$

Notice that (a), (b), and the first part of (c) are fulfilled. Property (d) follows from Lemma 4.4.2. To verify the second part of (c), write

$$B_\epsilon(z) = e^{i\beta}\prod_{p=1}^{n}\frac{w_p-z}{1-\overline{w_p}z},$$

in which $w_1, w_2, \ldots, w_n$ is a reindexing of (4.4.6). By direct calculation, we see that

$$\frac{B'_\epsilon(0)}{B_\epsilon(0)} = \sum_{p=1}^{n} \left(\overline{w_p} - \frac{1}{w_p} \right).$$

We have some freedom to control this expression. For example, if $w_1 = \epsilon e^{i\theta}$, then

$$\frac{B'_\epsilon(0)}{B_\epsilon(0)} = \left(\epsilon - \frac{1}{\epsilon} \right) e^{-i\theta} + \sum_{p=2}^{n} \left(\overline{w_p} - \frac{1}{w_p} \right).$$

If $B'_\epsilon(0) = 0$ for some choice of θ, we may change the value of θ so that, without violating the preceding properties, we obtain a B_ϵ such that $B'_\epsilon(0) \neq 0$. □

4.5 Generalized Rouché Theorem and Its Converse

Suppose that Γ is a simple closed curve in $\mathbb{C}$ (recall our standing convention that only piecewise C^1 curves are considered). We say that f *is analytic inside and on* Γ if there is a simply connected neighborhood of Γ upon which f is analytic. If f has no zeros on Γ, then the Argument Principle says that

$$Z_f(\Gamma) = \frac{1}{2\pi i} \int_\Gamma \frac{f'(z)}{f(z)}\, dz$$

is the number of zeros of f inside Γ, counted according to multiplicity. More generally, if f is meromorphic on and inside Γ and if f has no zeros or poles on Γ, then

$$Z_f(\Gamma) - P_f(\Gamma) = \frac{1}{2\pi i} \int_\Gamma \frac{f'(z)}{f(z)}\, dz,$$

in which $P_f(\Gamma)$ denotes the number of poles of f inside Γ.

Rouché's Theorem asserts that if f and h are analytic inside and on a simple closed curve Γ and if $|h| < |f|$ on Γ, then

$$Z_f(\Gamma) = Z_{f+h}(\Gamma).$$

Here is an intuitive explanation that can be found in many standard complex analysis texts. A person walks a dog around the origin along the path $f(z)$, in which $z \in \Gamma$. The dog's position is $f(z) + h(z)$, in which $z \in \Gamma$. The condition $|h| < |f|$ means that the leash is always shorter than the distance from the walker to the origin. Therefore, the walker and the dog circle the origin the same number of times (no matter how many times the dog circles the walker). Hence f and $f + h$ have the same number of zeros inside of Γ by the Argument Principle.

A stronger version of Rouché's theorem is the following. It implies the weaker form discussed above; see Exercise 4.9.

Theorem 4.5.1 (Glicksberg [64]) *If f and g are analytic inside and on a simple closed curve Γ and if*

$$|f - g| < |f| + |g| \text{ on } \Gamma, \tag{4.5.2}$$

then f and g have the same number of zeros inside Γ.

Proof First note that (4.5.2) implies that

$$\left|\frac{f}{g} - 1\right| < \left|\frac{f}{g}\right| + 1$$

on Γ. The strict inequality above tells us that $f/g \leqslant 0$ cannot occur anywhere on Γ. In particular, Γ does not pass through a zero of either f or g. Thus, f/g maps Γ into $\mathbb{C}\backslash(-\infty, 0]$. If $\log z$ is the principal branch of the logarithm, then

$$\frac{d}{dz}\log\left(\frac{f}{g}\right) = \frac{(f/g)'}{f/g}$$

on some open set containing Γ. Consequently,

$$\begin{aligned}
0 &= \frac{1}{2\pi}\int_\Gamma \frac{d}{dz}\log\left(\frac{f(z)}{g(z)}\right)dz \\
&= \frac{1}{2\pi i}\int_\Gamma \frac{(f(z)/g(z))'}{f(z)/g(z)}\, dz \\
&= \frac{1}{2\pi i}\int_\Gamma \left(\frac{f'(z)g(z) - f(z)g'(z)}{g(z)^2} \cdot \frac{g(z)}{f(z)}\right) dz \\
&= \frac{1}{2\pi i}\int_\Gamma \left(\frac{f'(z)}{f(z)} - \frac{g'(z)}{g(z)}\right) dz \\
&= \frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}\, dz - \frac{1}{2\pi i}\int_\Gamma \frac{g'(z)}{g(z)}\, dz \\
&= Z_f(\Gamma) - Z_g(\Gamma)
\end{aligned}$$

by the Argument Principle. □

Since we are primarily concerned with functions on the unit disk $\mathbb{D}$, by Z_f we mean $Z_f(\mathbb{T})$. That is, Z_f, when used without reference to a curve, denotes the number of zeros of f inside $\mathbb{D}$. Similarly, if P_f denotes the number of poles of f inside of $\mathbb{D}$. In all cases, we restrict our attention to functions that have no zeros or poles on $\mathbb{T}$.

Suppose that f and g are analytic on $|z| < R$ for some $R > 1$ and have no zeros on $\mathbb{T}$. Let B_1 and B_2 be finite Blaschke products of degree n. If

$$|B_1 f + B_2 g| < |f| + |g| \text{ on } \mathbb{T},$$

then

$$|B_1 f + B_2 g| < |B_1 f| + |B_2 g| \text{ on } \mathbb{T},$$

and hence Theorem 4.5.1 says that

$$n + Z_f = Z_{B_1 f} = Z_{B_2 g} = n + Z_g.$$

Consequently, f and g have the same number of zeros in $\mathbb{D}$. The converse of this result is also true.

Theorem 4.5.3 (Challener–Rubel [20]) *Suppose that f and g are analytic on $|z| < R$ for some $R > 1$ and that they have no zeros on $\mathbb{T}$. If f and g have the same number of zeros in $\mathbb{D}$, then there are finite Blaschke products B_1 and B_2 of the same degree such that*

$$|B_1 f + B_2 g| < |f| + |g| \textit{ on } \mathbb{T}.$$

Proof Since f and g are continuous with no zeros on $\mathbb{T}$,

$$m = \min_{\zeta \in \mathbb{T}} \min \{|f(\zeta)|, |g(\zeta)|\} > 0 \quad \text{and} \quad M = \max_{\zeta \in \mathbb{T}} \max \{|f(\zeta)|, |g(\zeta)|\} < \infty.$$

Let $h = g/f$ and $u = h/|h|$. Since $-u$ is a continuous unimodular function on $\mathbb{T}$, Theorem 4.3.1 provides two finite Blaschke products B_1 and B_2 so that

$$\max_{\xi \in \mathbb{T}} \left| u(\xi) + \frac{B_1(\xi)}{B_2(\xi)} \right| < \frac{m}{M}.$$

On $\mathbb{T}$, we have

$$\begin{aligned}
|B_1 f + B_2 g| &= |f||B_2| \left| \frac{B_1}{B_2} + \frac{g}{f} \right| \\
&= |f| \left| h + \frac{B_1}{B_2} \right| \\
&= |f| \left| u|h| + \frac{B_1}{B_2} \right| \\
&= |f| \left| u(|h| - 1) + \left(u + \frac{B_1}{B_2} \right) \right|
\end{aligned}$$

$$
\begin{aligned}
&\leqslant |f|\left|\frac{|g|}{|f|}-1\right|+|f|\left|u+\frac{B_1}{B_2}\right| \\
&< \big||g|-|f|\big|+M\cdot\frac{m}{M} \\
&= \big||f|-|g|\big|+m \\
&< |f|+|g|
\end{aligned}
$$

by the definitions of m and M; see Exercise 4.10. Since $|f| = |B_1 f|$ and $|g| = |B_2 g|$ on $\mathbb{T}$, the strong version of Rouché's Theorem (Theorem 4.5.1) shows that $B_1 f$ and $B_2 g$ have the same number of zeros inside $\mathbb{D}$. Since f and g have the same number of zeros inside $\mathbb{D}$, the finite Blaschke products B_1 and B_2 have the same degree. □

See Exercises 4.11 and 4.12 for meromorphic versions of the preceding theorem.

4.6 Exercises

4.1 Show that the function

$$f(z) = \exp\Big(-\frac{1+z}{1-z}\Big)$$

belongs to the Schur class but does not have a continuous extension to $\mathbb{D}^-$.

4.2 This exercise continues the discussion in Exercise 3.3 that compares finite Blaschke products and polynomials. Prove that given any entire function f there is a sequence of polynomials P_n that converges pointwise to f. The reader will recognize this as the analogue of Theorem 4.1.1.

4.3 Verify (4.1.3).

4.4 Show that if $f \in \mathscr{S}$ and $f(0) = 0$, then there is a sequence of finite Blaschke products B_n with $B_n(0) = 0$ for all n and such that $B_n \to f$ uniformly on compact subsets of $\mathbb{D}$.

4.5 Show that the product of two convex combinations of finite Blaschke products is a convex combination of finite Blaschke products.

4.6 Verify (4.2.3) and (4.2.4).

4.7 Generalize Theorem 4.4.1 by showing that we can choose B_ϵ such that

$$B_\epsilon^{(j)}(0) \neq 0, \qquad 0 \leqslant j \leqslant n-1.$$

4.8 Verify (4.4.5).

4.9 Prove that the strong version of Rouché's theorem (Theorem 4.5.1) implies the weak version of Rouché's theorem.

4.10 If f and g are continuous and nonzero on $\mathbb{T}$, prove that

$$\big||f| - |g|\big| + m \leqslant |f| + |g|,$$

in which

$$m = \min_{\zeta \in \mathbb{T}} \min \big\{|f(\zeta)|, |g(\zeta)|\big\} > 0.$$

4.11 This exercise concerns a meromorphic generalization of Theorem 4.5.3. Suppose that f and g are meromorphic on $|z| < R$ for some $R > 1$ and that they have no zeros or poles on $\mathbb{T}$. Then

$$Z_f - P_f = Z_g - P_g$$

if and only if there are finite Blaschke products B_1 and B_2 of the same degree such that

$$|B_1 f + B_2 g| < |f| + |g| \text{ on } \mathbb{T}.$$

4.12 This is a generalization of Exercise 4.11, in which we no longer assume that the functions involved are meromorphic on a neighborhood of $\mathbb{D}$. Suppose that f and g are continuous on $\mathbb{D}^-$, with the exception of finitely many poles in $\mathbb{D}$. Furthermore, suppose that f and g have no zeros on $\mathbb{T}$. Show that

$$Z_f - P_f = Z_g - P_g$$

if and only if there are finite Blaschke products B_1 and B_2 of the same degree such that

$$|B_1 f + B_2 g| < |f| + |g| \text{ on } \mathbb{T}.$$

Chapter 5
Zeros and Residues

5.1 Gauss–Lucas Theorem

There is a fascinating relationship between the zeros of a finite Blaschke product B and the location of the solutions of the equation $B(z) = \zeta$, in which $\zeta \in \mathbb{T}$ is fixed. There are also results concerning the relationship between the zeros of B and those of B' (discussed in the next chapter). To place all of these in context, we begin with an old theorem of Gauss and Lucas [99]. Recall that if $z_1, z_2, \ldots, z_n \in \mathbb{C}$, then

$$\operatorname{conv}\{z_1, z_2, \ldots, z_n\} = \left\{ \sum_{j=1}^{n} \lambda_j z_j : \lambda_j \in [0, 1], \sum_{j=1}^{n} \lambda_j = 1 \right\}$$

is the convex hull of the points $z_1, z_2, \ldots, z_n$. This next theorem is stronger than the corollary that follows it, which is more commonly known as the Gauss–Lucas theorem.

Theorem 5.1.1 (Gauss–Lucas) *Suppose $z_1, z_2, \ldots, z_n \in \mathbb{C}$ and $c_1, c_2, \ldots, c_n$ are positive. Then*

$$f(z) = \frac{c_1}{z - z_1} + \frac{c_2}{z - z_2} + \cdots + \frac{c_n}{z - z_n} \tag{5.1.2}$$

has at most $n - 1$ zeros, all of which belong to $\operatorname{conv}\{z_1, z_2, \ldots, z_n\}$.

Proof Multiply both sides of the equation $f(z) = 0$ by

$$(z - z_1)(z - z_2) \cdots (z - z_n)$$

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_5

and simplify to obtain a polynomial equation of degree at most $n-1$. Thus, f has at most $n-1$ zeros in $\mathbb{C}$, repeated according to multiplicity. If w is a zero of (5.1.2), then

$$\frac{c_1}{w-z_1} + \frac{c_2}{w-z_2} + \cdots + \frac{c_n}{w-z_n} = 0. \tag{5.1.3}$$

Since $c_1, c_2, \ldots, c_n \in \mathbb{R}$, (5.1.3) is equivalent to

$$\frac{c_1}{\overline{w}-\overline{z_1}} + \frac{c_2}{\overline{w}-\overline{z_2}} + \cdots + \frac{c_n}{\overline{w}-\overline{z_n}} = 0.$$

Multiply each summand in the previous equation by the appropriate

$$\frac{w-z_j}{w-z_j}$$

to conclude that (5.1.3) is equivalent to

$$\frac{c_1(w-z_1)}{|w-z_1|^2} + \frac{c_2(w-z_2)}{|w-z_2|^2} + \cdots + \frac{c_n(w-z_n)}{|w-z_n|^2} = 0.$$

Rewrite this as

$$\left(\frac{c_1}{|w-z_1|^2} + \cdots + \frac{c_n}{|w-z_n|^2}\right) w = \frac{c_1}{|w-z_1|^2} z_1 + \cdots + \frac{c_n}{|w-z_n|^2} z_n.$$

Thus,

$$w = \lambda_1 z_1 + \lambda_2 z_2 + \cdots + \lambda_n z_n,$$

in which

$$\lambda_j = \frac{\dfrac{c_j}{|w-z_j|^2}}{\dfrac{c_1}{|w-z_1|^2} + \cdots + \dfrac{c_n}{|w-z_n|^2}}, \qquad 1 \leqslant j \leqslant n.$$

Since $c_1, c_2, \ldots, c_n > 0$, it follows that

$$0 < \lambda_1, \lambda_2, \ldots, \lambda_n < 1 \quad \text{and} \quad \lambda_1 + \lambda_2 + \cdots + \lambda_n = 1.$$

Consequently, w lies in the convex hull of $\{z_1, z_2, \ldots, z_n\}$. □

The following corollary is itself sometimes referred to as the Gauss–Lucas theorem. See Exercises 6.1 and 6.2 for special cases.

Corollary 5.1.4 (Gauss–Lucas) *If P is a nonconstant polynomial, then the zeros of the derivative P' are contained in the convex hull of the zeros of P.*

Proof Without loss of generality, suppose that P is monic and write

$$P(z) = (z-z_1)(z-z_2)\cdots(z-z_n).$$

Then the zeros of the rational function

$$\frac{P'(z)}{P(z)} = \frac{1}{z - z_1} + \frac{1}{z - z_2} + \cdots + \frac{1}{z - z_n}$$

are the zeros of P'. Theorem 5.1.1 asserts that the $n - 1$ zeros of P', counted according to multiplicity, are contained in $\mathrm{conv}\{z_1, z_2, \ldots, z_n\}$. □

5.2 Gauss–Lucas Theorem for Finite Blaschke Products

The Gauss–Lucas theorem can be used to prove a beautiful result (Theorem 5.2.8 below) about the location of the zeros of a finite Blaschke product in terms of its boundary values [29, 66]. We first need some information about the derivative of a finite Blaschke product.

The following identity, which generalizes (1.6.1), was used by Frostman to discuss the boundary properties of the derivative of infinite Blaschke products [50] and by Pekarskiĭ [113] to estimate the derivative of a Cauchy transform.

Theorem 5.2.1 *Let*

$$B(z) = \prod_{j=1}^{n} \frac{z - z_j}{1 - \overline{z_j} z}, \tag{5.2.2}$$

$B_1 = 1$, *and*

$$B_k(z) = \prod_{j=1}^{k-1} \frac{z - z_j}{1 - \overline{z_j} z}, \quad 2 \leqslant k \leqslant n.$$

Then for each $z \in \mathbb{C} \backslash \mathbb{T}$,

$$\frac{1 - |B(z)|^2}{1 - |z|^2} = \sum_{k=1}^{n} |B_k(z)|^2 \frac{1 - |z_k|^2}{|1 - \overline{z_k} z|^2}. \tag{5.2.3}$$

Proof We proceed by induction on n. The case $n = 1$ is (1.6.1). Suppose that (5.2.3) holds for any finite Blaschke product of degree $n - 1$. By our inductive hypothesis,

$$\frac{1 - |B_n(z)|^2}{1 - |z|^2} = \sum_{k=1}^{n-1} |B_k(z)|^2 \frac{1 - |z_k|^2}{|1 - \overline{z_k} z|^2}. \tag{5.2.4}$$

Since

$$B(z) = B_n(z) \frac{z - z_n}{1 - \overline{z_n} z},$$

we have

$$
\begin{aligned}
1 - |B(z)|^2 &= 1 - |B_n(z)|^2 \left| \frac{z_n - z}{1 - \overline{z_n} z} \right|^2 \\
&= 1 - |B_n(z)|^2 + |B_n(z)|^2 \left(1 - \left| \frac{z_n - z}{1 - \overline{z_n} z} \right|^2 \right) \\
&= 1 - |B_n(z)|^2 + |B_n(z)|^2 \frac{(1 - |z|^2)(1 - |z_n|^2)}{|1 - \overline{z_n} z|^2}.
\end{aligned}
$$

Divide the previous expression by $1 - |z|^2$ to obtain

$$
\frac{1 - |B(z)|^2}{1 - |z|^2} = \frac{1 - |B_n(z)|^2}{1 - |z|^2} + |B_n(z)|^2 \frac{1 - |z_n|^2}{|1 - \overline{z_n} z|^2}
$$

and then use the inductive hypothesis (5.2.4) to see the preceding equals

$$
\sum_{k=1}^{n} |B_k(z)|^2 \frac{1 - |z_k|^2}{|1 - \overline{z_k} z|^2}.
$$

This completes the proof. □

Since B' is continuous on $\mathbb{D}^-$, we have

$$
\lim_{z \to \zeta} |B'(z)| = |B'(\zeta)|, \qquad \zeta \in \mathbb{T},
$$

in which the convergence is uniform with respect to ζ. An interesting expression for $|B'(\zeta)|$ is provided by the following theorem.

Theorem 5.2.5 *If B is a finite Blaschke product, then*

$$
\lim_{z \to \zeta} \frac{1 - |B(z)|^2}{1 - |z|^2} = \lim_{z \to \zeta} \frac{1 - |B(z)|}{1 - |z|} = |B'(\zeta)|, \qquad \zeta \in \mathbb{T},
$$

in which convergence is uniform with respect to ζ.

Proof Let B denote the finite Blaschke product (5.2.2). Recall (3.4.8), which says that

$$
|B'(\zeta)| = \sum_{k=1}^{n} \frac{1 - |z_k|^2}{|\zeta - z_k|^2}.
$$

From (5.2.3) we get

$$\begin{aligned}\lim_{z\to\zeta}\frac{1-|B(z)|^2}{1-|z|^2} &= \lim_{z\to\zeta}\sum_{k=1}^{n}|B_k(z)|^2\frac{1-|z_k|^2}{|1-\overline{z_k}z|^2}\\ &= \sum_{k=1}^{n}\frac{1-|z_k|^2}{|1-\overline{z_k}\zeta|^2}\\ &= \sum_{k=1}^{n}\frac{1-|z_k|^2}{|\zeta-z_k|^2}\\ &= |B'(\zeta)|. \end{aligned} \tag{5.2.6}$$

Since the modulus of a finite Blaschke product tends uniformly to 1 as one approaches the boundary $\mathbb{T}$, (5.2.3) shows that the convergence in (5.2.6) is uniform in ζ. Consequently,

$$\begin{aligned}\lim_{z\to\zeta}\frac{1-|B(z)|^2}{1-|z|^2} &= \lim_{z\to\zeta}\frac{1-|B(z)|}{1-|z|}\cdot\frac{1+|B(z)|}{1+|z|}\\ &= \lim_{z\to\zeta}\frac{1-|B(z)|}{1-|z|}. \quad\square\end{aligned}$$

The next lemma allows us to bring in the Gauss–Lucas theorem (Theorem 5.1.1). Before proceeding, recall that Theorem 3.4.10 ensures that for each $w\in\mathbb{T}$, the equation $B(z)=w$ has exactly n distinct solutions.

Lemma 5.2.7 *Let $z_1, z_2, \ldots, z_{n-1}\in\mathbb{D}$,*

$$B(z) = z\prod_{k=1}^{n-1}\frac{z_k-z}{1-\overline{z_k}z},$$

and $w\in\mathbb{T}$. Let $\zeta_1,\zeta_2,\ldots,\zeta_n\in\mathbb{T}$ be the n distinct solutions to $B(\zeta)=w$ and define

$$\lambda_k = \left(1+\sum_{j=1}^{n-1}\frac{1-|z_j|^2}{|\zeta_k-z_j|^2}\right)^{-1}, \qquad 1\leqslant k\leqslant n.$$

Then

$$0<\lambda_1,\lambda_2,\ldots,\lambda_n<1 \quad \text{and} \quad \lambda_1+\lambda_2+\cdots+\lambda_n = 1.$$

Moreover,

$$\frac{B(z)/z}{B(z)-w} = \frac{(z-z_1)\cdots(z-z_{n-1})}{(z-\zeta_1)\cdots(z-\zeta_n)} = \frac{\lambda_1}{z-\zeta_1}+\cdots+\frac{\lambda_n}{z-\zeta_n}.$$

Proof Observe that

$$\frac{B(z)/z}{B(z)-w} = \frac{P(z)}{Q(z)},$$

in which P and Q are polynomials with $\deg P = n-1$ and $\deg Q = n$. The zeros of P are $z_1, z_2, \ldots, z_{n-1}$ and the zeros of Q are $\zeta_1, \zeta_2, \ldots, \zeta_n$. Hence

$$\frac{B(z)/z}{B(z)-w} = \alpha \frac{(z-z_1)\cdots(z-z_{n-1})}{(z-\zeta_1)\cdots(z-\zeta_n)}$$

for some constant $\alpha \neq 0$. Multiply the preceding by z, let $z \to \infty$, and conclude that $\alpha = 1$. Now perform a partial fraction expansion and obtain

$$\frac{B(z)/z}{B(z)-w} = \frac{\lambda_1}{z-\zeta_1} + \cdots + \frac{\lambda_n}{z-\zeta_n}$$

for some $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$. Fix $k = 1, 2, \ldots, n$, multiply the previous identity by $z - \zeta_k$, and let $z \to \zeta_k$ to see that

$$\begin{aligned}
\lambda_k &= \lim_{z \to \zeta_k} \frac{B(z)}{z} \cdot \frac{z-\zeta_k}{B(z)-w} \\
&= \frac{B(\zeta_k)}{\zeta_k B'(\zeta_k)} \\
&= \frac{1}{1 + \displaystyle\sum_{j=1}^{n-1} \frac{1-|z_j|^2}{|\zeta_k - z_j|^2}} \qquad \text{(by (3.4.7))}.
\end{aligned}$$

Consequently, $0 < \lambda_k < 1$. Let $z \to \infty$ in the identity

$$\frac{B(z)}{B(z)-w} = \frac{\lambda_1 z}{z-\zeta_1} + \cdots + \frac{\lambda_n z}{z-\zeta_n}$$

and conclude that

$$\lambda_1 + \cdots + \lambda_n = \lim_{z \to \infty} \frac{B(z)}{B(z)-w} = 1. \qquad \square$$

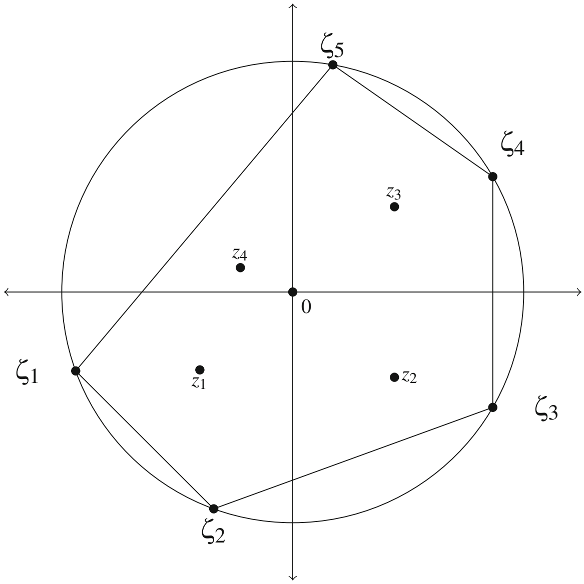
We are now ready to state and prove the main theorem of this section.

Theorem 5.2.8 *Let* $z_1, z_2, \ldots, z_{n-1} \in \mathbb{D}$,

$$B(z) = z \prod_{k=1}^{n-1} \frac{z_k - z}{1 - \overline{z_k} z},$$

$w \in \mathbb{T}$, *and let* $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ *be the* n *distinct solutions to* $B(\zeta) = w$. *Then* $z_1, z_2, \ldots, z_{n-1}$ *belong to* $\operatorname{conv}\{\zeta_1, \zeta_2, \ldots, \zeta_n\}$.

Fig. 5.1 conv$\{\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5\}$ contains z_1, z_2, z_3, z_4 (the origin is not necessarily contained in the convex hull)



Proof Lemma 5.2.7 yields the representation

$$\frac{B(z)/z}{B(z) - w} = \frac{\lambda_1}{z - \zeta_1} + \cdots + \frac{\lambda_n}{z - \zeta_n}, \tag{5.2.9}$$

in which the right-hand side is a convex combination of the functions

$$\frac{1}{z - \zeta_1}, \quad \frac{1}{z - \zeta_2}, \quad \ldots, \quad \frac{1}{z - \zeta_n}.$$

Since the zeros of the quotient are precisely $z_1, z_2, \ldots, z_{n-1}$, Theorem 5.1.1 says that they belong to conv$\{\zeta_1, \zeta_2, \ldots, \zeta_n\}$. □

Figure 5.1 illustrates Theorem 5.2.8 for a finite Blaschke product of degree $n = 5$.

Corollary 5.2.10 *Let* $z_1, z_2, \ldots, z_{n-1} \in \mathbb{D}$,

$$B(z) = z \prod_{k=1}^{n-1} \frac{z_k - z}{1 - \overline{z_k} z},$$

$w \in \mathbb{T}$, $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ *be the* n *distinct solutions to* $B(\zeta) = w$, *and*

$$\lambda_k = \left(1 + \sum_{j=1}^{n-1} \frac{1 - |z_j|^2}{|\zeta_k - z_j|^2}\right)^{-1}, \qquad 1 \leqslant k \leqslant n.$$

Then $\lambda_1, \lambda_2, \ldots, \lambda_n \in (0, 1)$, $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$, *and*

$$\frac{1}{1 - \overline{w}B(z)} = \frac{\lambda_1}{1 - \overline{\zeta}_1 z} + \cdots + \frac{\lambda_n}{1 - \overline{\zeta}_n z}. \tag{5.2.11}$$

Proof Lemma 5.2.7 yields

$$\frac{B(z)/z}{B(z) - w} = \frac{\lambda_1}{z - \zeta_1} + \cdots + \frac{\lambda_n}{z - \zeta_n}, \tag{5.2.12}$$

which is valid for all $z \in \mathbb{T}$ if properly interpreted at the poles. For such points, $z\overline{z} = 1$ and $B(z)\overline{B(z)} = 1$ and so we can write (5.2.12) as

$$\frac{1}{1 - w\overline{B(z)}} = \frac{\lambda_1}{1 - \zeta_1\overline{z}} + \cdots + \frac{\lambda_n}{1 - \zeta_n\overline{z}}, \qquad z \in \mathbb{T}.$$

Conjugate the preceding and obtain (5.2.11) for $z \in \mathbb{T}$. Since both sides of (5.2.11) are meromorphic on $\mathbb{C}$, the identity holds everywhere. □

5.3 Zeros as Foci of an Ellipse

For the finite Blaschke product

$$B(z) = z\left(\frac{\alpha - z}{1 - \overline{\alpha}z}\right), \qquad \alpha \in \mathbb{D}\backslash\{0\},$$

Theorem 5.2.8 has an interesting geometric interpretation. Any line that passes through α intersects $\mathbb{T}$ at two distinct points ζ_1 and ζ_2. According to Theorem 5.2.8, $B(\zeta_1) = B(\zeta_2)$. Conversely, if $\zeta_1, \zeta_2 \in \mathbb{T}$ are such that $B(\zeta_1) = B(\zeta_2)$, then α must be on the line that connects ζ_1 to ζ_2; Fig. 5.2.

In Theorem 5.2.8, we are free to choose any $w \in \mathbb{T}$ and obtain the n distinct solutions $\zeta_{1,w}, \zeta_{2,w}, \ldots, \zeta_{n,w}$ to $B(\zeta) = w$. Therefore, the points $z_1, z_2, \ldots, z_{n-1}$ are contained in

$$\bigcap_{w \in \mathbb{T}} \operatorname{conv}\{\zeta_{1,w}, \zeta_{2,w}, \ldots, \zeta_{n,w}\} \tag{5.3.1}$$

For a finite Blaschke product of degree three, this phenomenon is depicted in Fig. 5.3. It appears as if the intersection (5.3.1) determines an ellipse (Fig. 5.4). This is not a coincidence.

Theorem 5.3.2 (Daepp–Gorkin–Mortini [29]) *Let*

$$B(z) = z\left(\frac{z_1 - z}{1 - \overline{z_1}z}\right)\left(\frac{z_2 - z}{1 - \overline{z_2}z}\right),$$

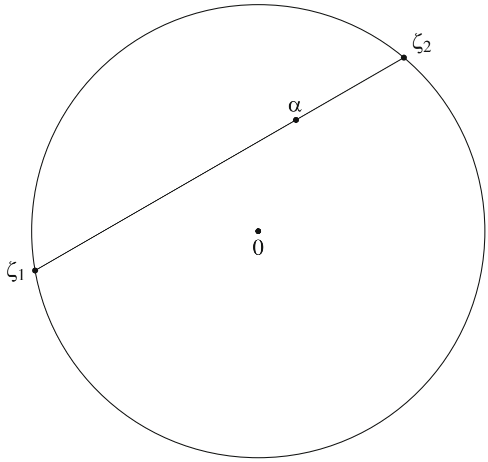


Fig. 5.2 Two points ζ_1, ζ_2 with the same image under B

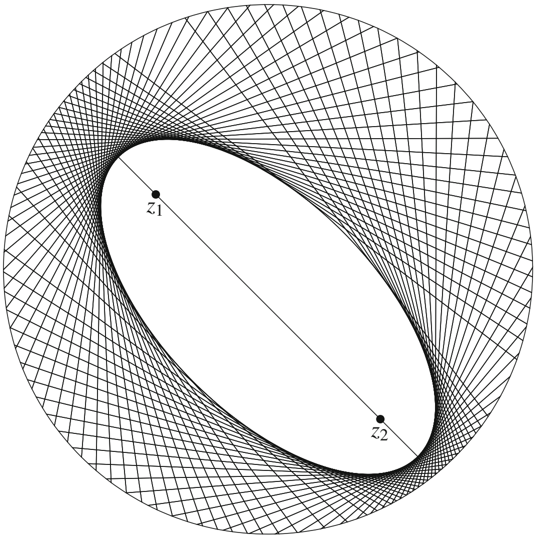


Fig. 5.3 $\bigcap_{w\in\mathbb{T}} \text{conv}\{\zeta_{1,w}, \zeta_{2,w}, \zeta_{3,w}\}$ for a Blaschke product of degree three that vanishes at the origin (the origin is not necessarily contained in the intersection)

in which $z_1, z_2 \in \mathbb{D}\backslash\{0\}$ are distinct. Let $w \in \mathbb{T}$ and let $\zeta_1, \zeta_2, \zeta_3$ be the distinct solutions of $B(\zeta) = w$. Let

$$\lambda_j = \left(1 + \frac{1-|z_1|^2}{|\zeta_j - z_1|^2} + \frac{1-|z_2|^2}{|\zeta_j - z_2|^2}\right)^{-1}, \qquad j = 1, 2, 3. \tag{5.3.3}$$

Then the line that passes through ζ_1 and ζ_2 is tangent to the ellipse

$$E = \{z : |z - z_1| + |z - z_2| = |1 - \overline{z_1} z_2|\} \tag{5.3.4}$$

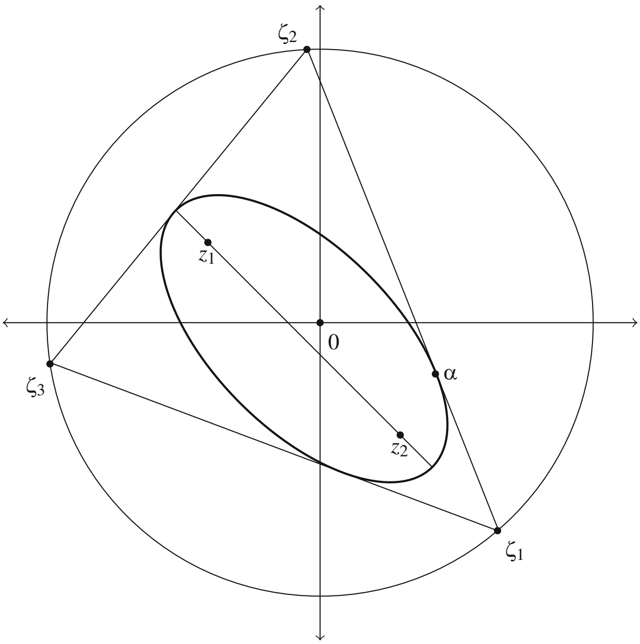


Fig. 5.4 The line through ζ_1 and ζ_2 is tangent to the ellipse E from (5.3.4)

at the point

$$\alpha = \frac{\lambda_2}{\lambda_1 + \lambda_2}\,\zeta_1 + \frac{\lambda_1}{\lambda_1 + \lambda_2}\,\zeta_2.$$

Conversely, each tangent line of E intersects $\mathbb{T}$ at two points η_1, η_2 such that $B(\eta_1) = B(\eta_2)$.

Proof Fix $w \in \mathbb{T}$. Lemma 5.2.7 ensures that

$$\begin{aligned}\frac{B(z)/z}{B(z) - w} &= \frac{(z - z_1)(z - z_2)}{(z - \zeta_1)(z - \zeta_2)(z - \zeta_3)} \\ &= \frac{\lambda_1}{z - \zeta_1} + \frac{\lambda_2}{z - \zeta_2} + \frac{\lambda_3}{z - \zeta_3},\end{aligned}$$

in which $\lambda_1, \lambda_2, \lambda_3$ are given by (5.3.3); in particular, $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Since

$$\lambda_1 + \lambda_2 = 1 - \lambda_3,$$

we have

$$\frac{B(z)/z}{B(z) - w} = \frac{(z - z_1)(z - z_2)}{(z - \zeta_1)(z - \zeta_2)(z - \zeta_3)} = \frac{(1 - \lambda_3)(z - \alpha)}{(z - \zeta_1)(z - \zeta_2)} + \frac{\lambda_3}{z - \zeta_3}; \qquad (5.3.5)$$

see Exercise 5.1. Set $z = z_1$ and then $z = z_2$ to obtain

$$\frac{z_1 - \alpha}{(z_1 - \zeta_1)(z_1 - \zeta_2)} + \frac{m}{z_1 - \zeta_3} = \frac{z_2 - \alpha}{(z_2 - \zeta_1)(z_2 - \zeta_2)} + \frac{m}{z_2 - \zeta_3} = 0, \tag{5.3.6}$$

in which $m = \lambda_3/(1 - \lambda_3)$. The first identity in (5.3.5) implies that

$$B(z) - w = \frac{(z - \zeta_1)\,(z - \zeta_2)\,(z - \zeta_3)}{(1 - \overline{z}_1\, z)\,(1 - \overline{z}_2\, z)}.$$

Substitute $z = z_1$ and $z = z_2$ into the preceding and obtain

$$\begin{aligned}\left|\frac{(z_1 - \zeta_1)(z_1 - \zeta_2)(z_1 - \zeta_3)}{(1 - |z_1|^2)(1 - \overline{z_2} z_1)}\right| &= \left|\frac{(z_2 - \zeta_1)(z_2 - \zeta_2)(z_2 - \zeta_3)}{(1 - \overline{z_2} z_1)(1 - |z_2|^2)}\right| \\ &= |w| \\ &= 1.\end{aligned}$$

Therefore,

$$\begin{aligned}|\alpha - z_1| + |\alpha - z_2| &= \frac{m|(z_1 - \zeta_1)(z_1 - \zeta_2)|}{|z_1 - \zeta_3|} + \frac{m|(z_2 - \zeta_1)(z_2 - \zeta_2)|}{|z_2 - \zeta_3|} \\ &= m|1 - \overline{z_2} z_1| \left(\frac{1 - |z_1|^2}{|z_1 - \zeta_3|^2} + \frac{1 - |z_2|^2}{|z_2 - \zeta_3|^2}\right) \\ &= m|1 - \overline{z_2} z_1| \left(\frac{1}{\lambda_3} - 1\right) \\ &= |1 - \overline{z_2} z_1|\end{aligned}$$

and hence $\alpha \in E$.

Set $z = \alpha$ in (5.3.5) and obtain

$$\frac{(\alpha - z_1)(\alpha - z_2)}{(\alpha - \zeta_1)(\alpha - \zeta_2)(\alpha - \zeta_3)} = \frac{\lambda_3}{\alpha - \zeta_3}.$$

Hence

$$\frac{z_1 - \alpha}{\zeta_1 - \alpha} \cdot \frac{z_2 - \alpha}{\zeta_2 - \alpha} = \lambda_3,$$

which implies that

$$\arg\left(\frac{z_1 - \alpha}{\zeta_1 - \alpha}\right) + \arg\left(\frac{z_2 - \alpha}{\zeta_2 - \alpha}\right) = 0.$$

In other words,

$$\angle \zeta_1 \alpha z_1 + \angle \zeta_2 \alpha z_2 = 0.$$

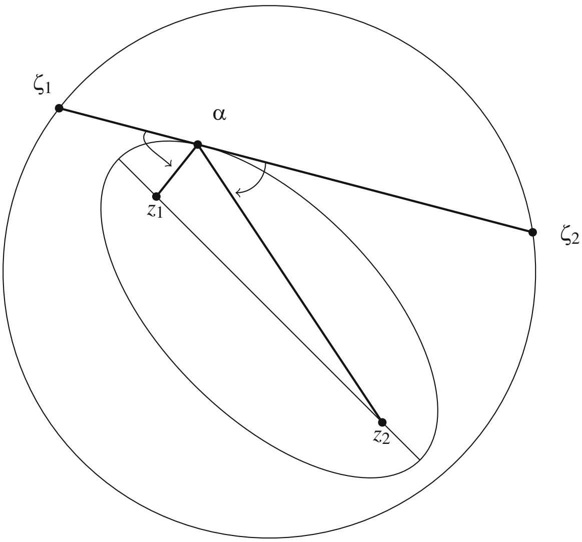


Fig. 5.5 The angles $\angle\zeta_1\alpha z_1$ and $\angle\zeta_2\alpha z_2$

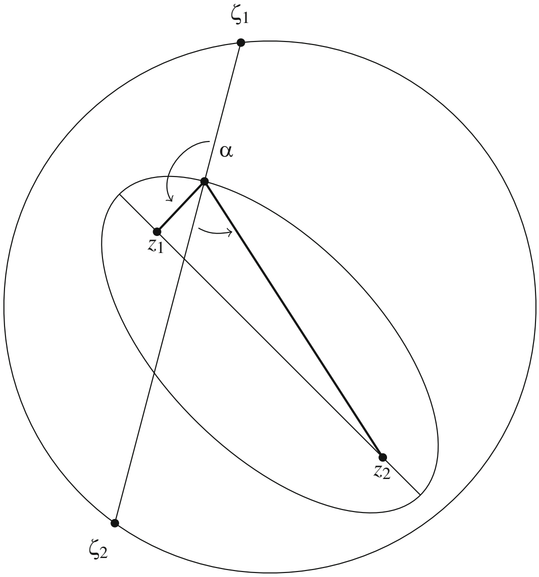


Fig. 5.6 The angles $\angle\zeta_1\alpha z_1$ and $\angle\zeta_2\alpha z_2$

Therefore, the line segment that connects ζ_1 and ζ_2 lies outside the triangle with vertices α, z_1, z_2 and makes equal angles with the sides of triangle at the vertex α; see Fig. 5.5. Based on well-known geometric properties of an ellipse, this property is fulfilled only by the tangent line at α.

To ensure that the line segment $\zeta_1\alpha\zeta_2$ is tangent to the ellipse, observe that this line cannot be the bisector of the angle $\angle z_1\alpha z_2$; see Fig. 5.6.

The last statement of the theorem is a consequence of the first part. From the point η_1 there are just two tangents to the ellipse E. Each of these tangent lines intersects $\mathbb{T}$ at one other point. One of these points is η_2. Call the other one η_3. Set $w = B(\eta_1)$. Based on the discussion above, these points must be the three distinct solutions of the equation $B(\zeta) = w$. □

5.4 A Weak Version of Sendov's Conjecture

There are various results about the relationship between the zeros of a polynomial and its derivative. The Gauss–Lucas theorem (Corollary 5.1.4) is a classic example from the vast literature on the subject. A famous conjecture in the area is due to Sendov.

Conjecture 5.4.1 (Sendov) If all the zeros $z_1, z_2, \dots, z_n$ of a polynomial P lie in $\mathbb{D}^-$, then each closed disk

$$D(z_k, 1)^- = \{z \in \mathbb{C} : |z - z_k| \leqslant 1\}, \qquad 1 \leqslant k \leqslant n,$$

contains at least one zero of P'.

Figure 5.7a illustrates a typical disk $D(z_k, 1)^-$ relevant to Sendov's conjecture. It turns out that the conjecture, in a stronger form, is true if one is given that $z_1, z_2, \dots, z_n \in \mathbb{T}$. In this case, each closed disk

$$D(\tfrac{z_k}{2}, \tfrac{1}{2})^- = \left\{z \in \mathbb{C} : |z - \tfrac{z_k}{2}| \leqslant \tfrac{1}{2}\right\}, \qquad 1 \leqslant k \leqslant n,$$

contains at least one zero of P' [124]. Since

$$D(\tfrac{z_k}{2}, \tfrac{1}{2})^- \subseteq D(z_k, 1)^-,$$

this result is stronger than what one expects from Sendov's conjecture alone; see Fig. 5.7b for an illustration of the preceding containment.

To see the relevance of Sendov's conjecture to the zeros of a finite Blaschke product, consider

$$B(z) = z \prod_{k=1}^{n-1} \frac{z_k - z}{1 - \overline{z}_k z},$$

in which $z_1, z_2, \dots, z_{n-1}$ are distinct points in $\mathbb{D}$. For each $w \in \mathbb{T}$, Theorem 3.4.10 asserts that the equation $B(z) = w$ has n distinct solutions $\zeta_1, \zeta_2, \dots, \zeta_n \in \mathbb{T}$. Theorem 5.2.8 tells us that $z_1, z_2, \dots, z_{n-1}$ belong to the convex hull of $\zeta_1, \zeta_2, \dots, \zeta_n$. We also established (5.2.9), which implies that

$$\frac{B(z)/z}{B(z) - w} = \frac{\lambda_1}{z - \zeta_1} + \cdots + \frac{\lambda_n}{z - \zeta_n} = \frac{P'(z)}{P(z)}, \tag{5.4.2}$$

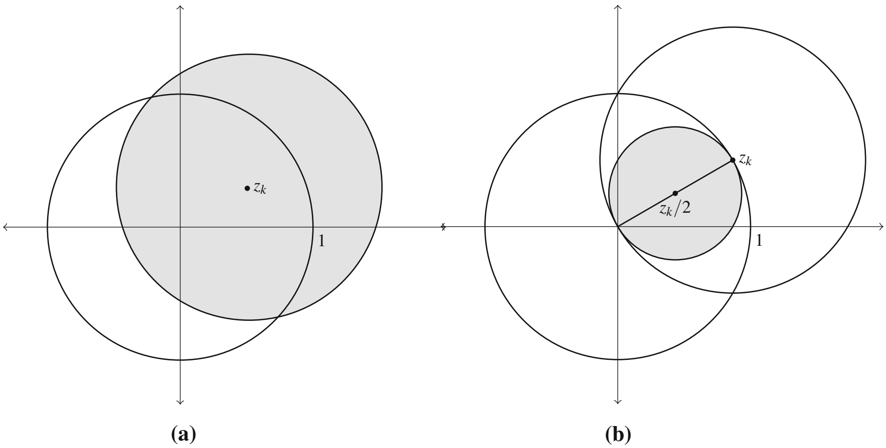


Fig. 5.7 Disks relevant to Sendov's conjecture. (**a**) The closed disk $D(z_k, 1)^-$. (**b**) The closed disk $D(\frac{z_k}{2}, \frac{1}{2})^-$

in which

$$P(z) = (z - \zeta_1)^{\lambda_1} \cdots (z - \zeta_n)^{\lambda_n}. \tag{5.4.3}$$

The fact that $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ permits us to define the preceding polynomial-like expression on $\mathbb{D}$ by using suitable branch cuts for each factor. We may think of $\zeta_1, \zeta_2, \ldots, \zeta_n$ as the zeros of P and $z_1, z_2, \ldots, z_{n-1}$ as the zeros of P'. Although P is not a polynomial, if each λ_i is rational, we can multiply both sides of (5.4.2) by an appropriate integer and then take P to be a polynomial. Therefore, we are naturally motivated to consider Sendov's conjecture in this case.

It is known that Sendov's conjecture does not hold for functions of the form (5.4.3); a simple counterexample is

$$P(z) = (1 + z)^{\epsilon/2}(1 - z)^{1-\epsilon/2},$$

in which $0 < \epsilon < 2$ [65]. The only zero of P' is $\epsilon - 1$. Thus, for $\epsilon \in (0, 2)$ and $\epsilon \neq 1$, one of the closed disks

$$D(1, 1)^- = \{z : |z - 1| \leqslant 1\} \quad \text{or} \quad D(-1, 1)^- = \{z : |z + 1| \leqslant 1\}$$

does not contain a zero of P'. However, we show that a weaker version of Sendov's conjecture holds for this family. This version reveals further the relationship between the locations of $\zeta_1, \zeta_2, \ldots, \zeta_n$ and $z_1, z_2, \ldots, z_{n-1}$. We require two simple facts that are needed in the proof of an important lemma below.

First, the Möbius transformation

$$\varphi(z) = \frac{1}{1 - z}$$

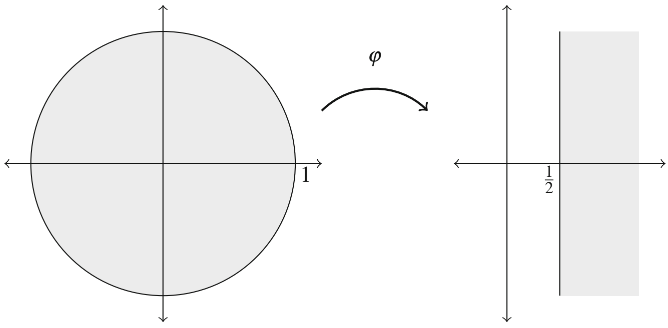


Fig. 5.8 The conformal mapping φ

maps the open unit disk bijectively onto the half plane $\{z \in \mathbb{C} : \operatorname{Re} z > 1/2\}$ and maps $\mathbb{T}\backslash\{1\}$ bijectively onto the line $\operatorname{Re} z = 1/2$. Hence,

$$z \in \mathbb{D}^{-} \quad \Longrightarrow \quad \operatorname{Re}\left(\frac{1}{1-z}\right) \geqslant \frac{1}{2}, \tag{5.4.4}$$

with equality if and only if $z \in \mathbb{T}$; see Fig. 5.8.

Second, if $z \in \mathbb{C}$, $\zeta \in \mathbb{T}$, and $r > 0$, then

$$\begin{aligned}\left|z - \frac{2r-1}{2r}\zeta\right|^2 &= \left|(1 - z\overline{\zeta}) - \frac{1}{2r}\right|^2 \\ &= |1 - z\overline{\zeta}|^2 - \frac{1}{r}\operatorname{Re}(1 - z\overline{\zeta}) + \frac{1}{4r^2} \\ &= |1 - z\overline{\zeta}|^2\left(1 - \frac{1}{r}\operatorname{Re}\left(\frac{1 - z\overline{\zeta}}{|1 - z\overline{\zeta}|^2}\right)\right) + \frac{1}{4r^2} \\ &= |1 - z\overline{\zeta}|^2\left(1 - \frac{1}{r}\operatorname{Re}\left(\frac{1}{1 - z\overline{\zeta}}\right)\right) + \frac{1}{4r^2}\end{aligned}$$

and hence

$$\operatorname{Re}\left(\frac{1}{1 - z\overline{\zeta}}\right) \geqslant r \quad \Longrightarrow \quad \left|z - \frac{2r-1}{2r}\zeta\right| \leqslant \frac{1}{2r}. \tag{5.4.5}$$

Theorem 5.4.6 (Gorkin–Rhoades [65]) *Suppose that $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ are distinct and $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$ are such that $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$. Let*

$$r_k = \frac{(n-1)\lambda_k}{(n-1)\lambda_k + (1 - \lambda_k)}, \qquad 1 \leqslant k \leqslant n, \tag{5.4.7}$$

and define

$$f(z) = \frac{\lambda_1}{z - \zeta_1} + \cdots + \frac{\lambda_n}{z - \zeta_n}.$$

Then for each $1 \leqslant k \leqslant n$, *there is a zero of* f *in the closed disk*

$$D((1 - r_k)\zeta_k, r_k)^-.$$

Proof Theorem 5.1.1 ensures that the zeros of f are in $\operatorname{conv}\{\zeta_1, \zeta_2, \ldots, \zeta_n\}$. Hence, all the zeros of f are in $\mathbb{D}^-$. In fact, the zeros of f are the roots of the polynomial

$$Q(z) = \sum_{k=1}^{n} \lambda_k \prod_{\substack{j=1 \\ j \neq k}}^{n} (z - \zeta_j).$$

Since $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$, the polynomial Q is monic and hence

$$Q(z) = \prod_{k=1}^{n-1} (z - z_k).$$

Taking the logarithmic derivative of the last two identities gives

$$\sum_{k=1}^{n-1} \frac{1}{z - z_k} = \frac{Q'(z)}{Q(z)} = \frac{\displaystyle\sum_{k=1}^{n} \lambda_k \sum_{\substack{j=1 \\ j \neq k}}^{n} \prod_{\substack{i=1 \\ i \neq j,k}}^{n} (z - \zeta_i)}{\displaystyle\sum_{k=1}^{n} \lambda_k \prod_{\substack{j=1 \\ j \neq k}}^{n} (z - \zeta_j)}.$$

Fix an $m \in \{1, 2, \ldots, n\}$, evaluate the preceding identity at $z = \zeta_m$, and obtain

$$\sum_{k=1}^{n-1} \frac{1}{\zeta_m - z_k} = \frac{\displaystyle\sum_{\substack{k=1 \\ k \neq m}}^{n} (\lambda_k + \lambda_m) \prod_{\substack{j=1 \\ j \neq k,m}}^{n} (\zeta_m - \zeta_j)}{\displaystyle\lambda_m \prod_{\substack{j=1 \\ j \neq m}}^{n} (\zeta_m - \zeta_j)}$$

$$= \sum_{\substack{k=1 \\ k \neq m}}^{n} \frac{(\lambda_k + \lambda_m)/\lambda_m}{\zeta_m - \zeta_k}.$$

Then (5.4.4) implies that

$$\sum_{k=1}^{n-1} \operatorname{Re}\left(\frac{1}{1 - z_k\overline{\zeta}_m}\right) = \sum_{\substack{k=1\\k\neq m}}^{n} \frac{\lambda_k + \lambda_m}{\lambda_m} \operatorname{Re}\left(\frac{1}{1 - \zeta_k\overline{\zeta}_m}\right)$$
$$= \sum_{\substack{k=1\\k\neq m}}^{n} \frac{\lambda_k + \lambda_m}{2\lambda_m}$$
$$= \frac{1 - \lambda_m}{2\lambda_m} + \frac{n-1}{2}.$$

The relation (5.4.4) also says that

$$\operatorname{Re}\left(\frac{1}{1 - z_k\overline{\zeta_m}}\right) \geqslant \frac{1}{2}$$

for all $1 \leqslant k \leqslant n-1$. Let k_0 be such that

$$\operatorname{Re}\left(\frac{1}{1 - z_{k_0}\overline{\zeta}_m}\right) = \max_{1\leqslant k\leqslant n-1} \operatorname{Re}\left(\frac{1}{1 - z_k\overline{\zeta}_m}\right).$$

Thus,

$$\operatorname{Re}\left(\frac{1}{1 - z_{k_0}\overline{\zeta}_m}\right) \geqslant \frac{1 - \lambda_m}{2(n-1)\lambda_m} + \frac{1}{2}.$$

Finally, apply (5.4.5) with $z = z_{k_0}$, $\zeta = \zeta_m$, and

$$r = \frac{1 - \lambda_m}{2(n-1)\lambda_m} + \frac{1}{2},$$

to obtain

$$|z_{k_0} - (1 - r_m)\zeta_m| \leqslant r_m. \qquad \square$$

Note that

$$D\big((1 - r_k)\zeta_k, r_k\big)^{-} \subseteq D(\zeta_k, 2r_k)^{-};$$

see Fig. 5.9. Thus, under the conditions of Theorem 5.4.6, there is a zero of f in the closed disk $D(\zeta_k, 2r_k)^{-}$.

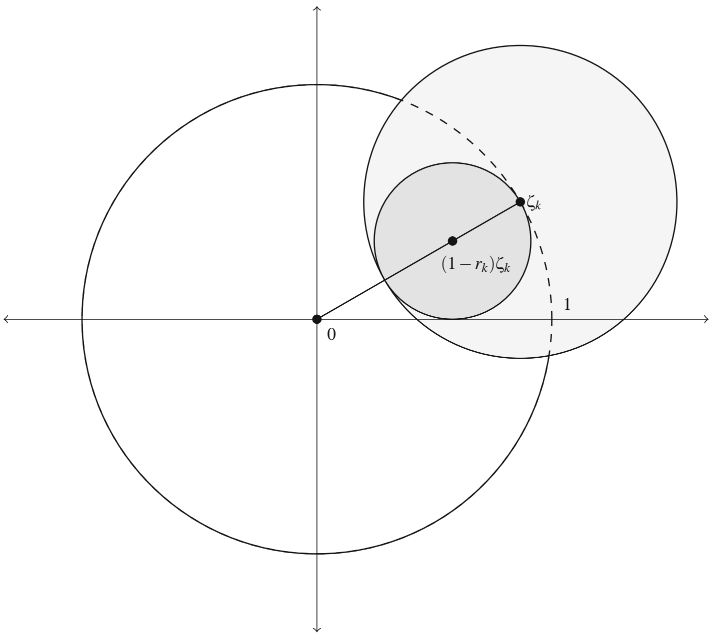


Fig. 5.9 The closed disks $D\big((1-r_k)\zeta_k, r_k\big)^-$ and $D(\zeta_k, 2r_k)^-$

5.5 A Forbidden Region

Theorems 5.1.1 and 5.4.6 tell us where the zeros of a function of the form

$$f(z) = \frac{\lambda_1}{z-\zeta_1} + \frac{\lambda_2}{z-\zeta_2} + \cdots + \frac{\lambda_n}{z-\zeta_n}, \tag{5.5.1}$$

in which

$$\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{C},$$

$$\lambda_1, \lambda_2, \ldots, \lambda_n > 0,$$

and

$$\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1,$$

might be; see Fig. 5.10. We now identify a region that excludes the zeros of f.

Theorem 5.5.2 (Gorkin–Rhoades [65]) *Let* $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ *be distinct and let* $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$ *be such that* $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$. *Let*

$$r = \min_{\substack{1 \leqslant i,j \leqslant n \\ i \neq j}} |\zeta_i - \zeta_j|$$

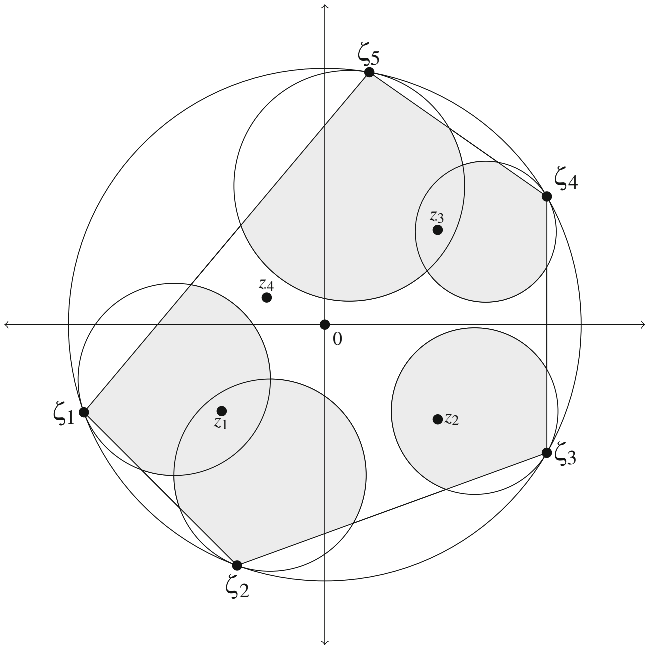


Fig. 5.10 The locus of the zeros of a finite Blaschke product of degree 5

and define

$$f(z) = \frac{\lambda_1}{z - \zeta_1} + \cdots + \frac{\lambda_n}{z - \zeta_n}.$$

Then for each $1 \leqslant k \leqslant n$*, there is no zero of* f *in the open disk* $D(\zeta_k, r\lambda_k)$.

Proof If z_0 is a zero of f, then

$$\frac{\lambda_1}{z_0 - \zeta_1} + \cdots + \frac{\lambda_n}{z_0 - \zeta_n} = 0.$$

For $1 \leqslant k \leqslant n$, the triangle inequality implies that

$$\frac{\lambda_k}{|z_0 - \zeta_k|} \leqslant \sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{\lambda_j}{|z_0 - \zeta_j|}. \tag{5.5.3}$$

However, for $j \neq k$

$$|z_0 - \zeta_j| \geqslant |\zeta_j - \zeta_k| - |z_0 - \zeta_k| \geqslant r - |z_0 - \zeta_k|.$$

If $|z_0 - \zeta_k| \geqslant r$, then $|z_0 - \zeta_k| \geqslant r\lambda_k$ and we are done. If $|z_0 - \zeta_k| < r$, then (5.5.3) implies that

$$\frac{\lambda_k}{|z_0 - \zeta_k|} \leqslant \frac{\sum_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j}{r - |z_0 - \zeta_k|} = \frac{1 - \lambda_k}{r - |z_0 - \zeta_k|},$$

which, after rearranging the terms, implies that $|z_0 - \zeta_k| \geqslant r\lambda_k$. □

5.6 The Best Citadel

The following result is a direct consequence of Theorems 5.4.6 and 5.5.2.

Theorem 5.6.1 (Gorkin–Rhoades [65]) *Let*

$$B(z) = z \prod_{k=1}^{n-1} \frac{z_k - z}{1 - \overline{z}_k\, z},$$

in which $z_1, z_2, \ldots, z_{n-1}$ are distinct points in $\mathbb{D}$, let $w \in \mathbb{T}$, and let $\zeta_1, \zeta_2, \ldots, \zeta_n$ denote the n distinct solutions of $B(\zeta) = w$. Let

$$r_k = \frac{(n-1)}{(n-1) + \displaystyle\sum_{j=1}^{n-1} \frac{1 - |z_j|^2}{|\zeta_k - z_j|^2}}, \qquad 1 \leqslant k \leqslant n,$$

and

$$\ell = \min_{\substack{1\leqslant i,j\leqslant n \\ i\neq j}} |\zeta_i - \zeta_j|.$$

For each $1 \leqslant k \leqslant n$, at least one z_i belongs to the closed disk

$$D((1 - r_k)\zeta_k, r_k)^-.$$

Moreover, none of the z_i belong to the open disk

$$D(\zeta_k, \ell\lambda_k).$$

Proof By Lemma 5.2.7,

$$\frac{B(z)/z}{B(z) - w} = \frac{(z - z_1)\cdots(z - z_{n-1})}{(z - \zeta_1)\cdots(z - \zeta_n)} = \frac{\lambda_1}{z - \zeta_1} + \cdots + \frac{\lambda_n}{z - \zeta_n},$$

in which

$$\lambda_k = \frac{1}{1 + \displaystyle\sum_{j=1}^{n-1} \frac{1 - |z_j|^2}{|\zeta_k - z_j|^2}}, \qquad 1 \leqslant k \leqslant n.$$

Observe that

$$0 < \lambda_1, \lambda_2, \ldots, \lambda_n < 1 \qquad \text{and} \qquad \lambda_1 + \lambda_2 + \cdots + \lambda_n = 1.$$

We are ready to apply Theorems 5.4.6 and 5.5.2. Via (5.4.7), the radius r_k is given by

$$\begin{aligned} r_k &= \frac{(n-1)\lambda_k}{(n-1)\lambda_k + (1-\lambda_k)} \\ &= \frac{(n-1)}{(n-1) + (1/\lambda_k - 1)} \\ &= \frac{(n-1)}{(n-1) + \displaystyle\sum_{j=1}^{n-1} \frac{1 - |z_j|^2}{|\zeta_k - z_j|^2}}. \end{aligned}$$

Theorem 5.4.6 implies that there is a z_i in the closed disk $D((1-r_k)\zeta_k, r_k)^-$. By Theorem 5.5.2, no z_i belongs to the open disk $D(\zeta_k, \ell\lambda_k)$. □

If we put Theorems 5.2.8 and 5.6.1 together, we get a better picture of the possible locations of the zeros of a finite Blaschke product; see Fig. 5.11.

5.7 Existence of a Nonzero Residue

The only entire finite Blaschke products are the unimodular scalar multiples of the monomials $1, z, z^2, \ldots$. All other finite Blaschke products have poles in $\mathbb{C}\backslash\mathbb{D}^-$ and hence we may consider their residues.

Theorem 5.7.1 (Heins [72]) *If B is a finite Blaschke product with at least one pole in $\mathbb{C}\backslash\mathbb{D}^-$, then B has a nonzero residue at some pole in $\mathbb{C}\backslash\mathbb{D}^-$.*

Proof Let

$$B(z) = e^{i\beta} z^m \prod_{n=1}^{N} \left(\frac{z - z_n}{1 - \overline{z_n} z} \right)^{m_n}, \tag{5.7.2}$$

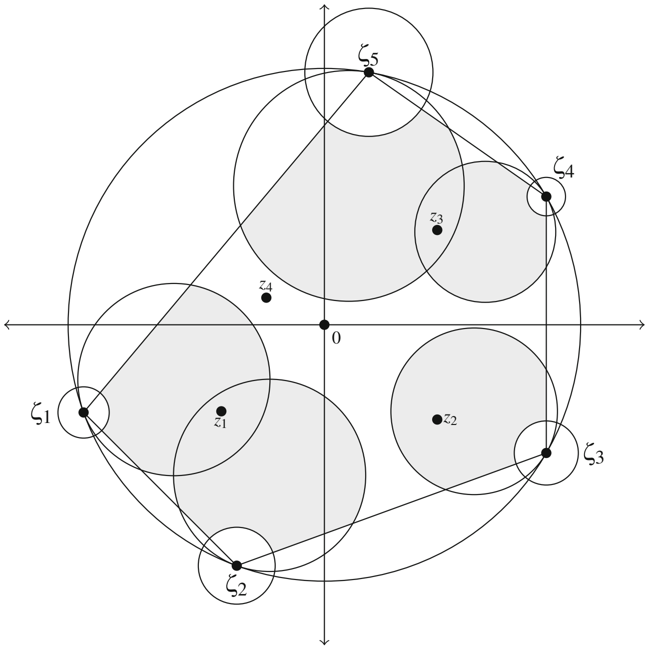


Fig. 5.11 The possible locus of zeros of B

in which $z_1, z_2, \ldots, z_n$ are the distinct zeros of B and let

$$\mathfrak{B}(z) = \int_0^z B(\zeta)\, d\zeta. \tag{5.7.3}$$

Since B is analytic on $\mathbb{D}^-$, the integral in (5.7.3) is independent of the path of integration. The Fundamental Theorem of Calculus says that $\mathfrak{B}'(z) = B(z)$ for each $z \in \mathbb{D}^-$ and that $\mathfrak{B}(0) = 0$. By (5.7.2), for each $e^{i\theta} \in \mathbb{T}$, we parameterize the straight line path from 0 to $e^{i\theta}$ by $r \mapsto re^{i\theta}$ for $0 \leqslant r \leqslant 1$, and obtain

$$\mathfrak{B}(e^{i\theta}) = \int_0^{e^{i\theta}} B(z)\, dz = \int_0^1 e^{i\beta} r^m e^{im\theta} \prod_{n=1}^{N} \left(\frac{re^{i\theta} - z_n}{1 - \overline{z_n} re^{i\theta}} \right)^{m_n} e^{i\theta} dr. \tag{5.7.4}$$

As a function of $z = re^{i\theta}$,

$$\prod_{n=1}^{N} \left(\frac{re^{i\theta} - z_n}{1 - \overline{z_n} re^{i\theta}} \right)^{m_n}$$

is a finite Blaschke product and hence (5.7.4) yields

$$
\begin{aligned}
|\mathfrak{B}(e^{i\theta})| &= \left| \int_0^1 e^{i\beta} r^m e^{im\theta} \prod_{n=1}^{N} \left(\frac{re^{i\theta} - z_n}{1 - \overline{z_n} r e^{i\theta}} \right)^{m_n} e^{i\theta} dr \right| \\
&\leqslant \int_0^1 r^m dr \\
&= \frac{1}{m+1}.
\end{aligned} \tag{5.7.5}
$$

Perform a partial fraction expansion on (5.7.2) to obtain

$$
B(z) = \sum_{n=0}^{m} \alpha_n z^n + \sum_{n=1}^{N} \sum_{\ell=1}^{m_n} \frac{\beta_{n,\ell}}{(1 - \overline{z_n} z)^{\ell}}. \tag{5.7.6}
$$

Suppose toward a contradiction that all of the residues of B are zero; that is

$$
\beta_{1,1} = \beta_{2,1} = \cdots = \beta_{N,1} = 0.
$$

By integration,

$$
\mathfrak{B}(z) = \sum_{n=0}^{m} \frac{\alpha_n}{n+1} z^{n+1} + \alpha + \sum_{n=1}^{N} \sum_{\ell=2}^{m_n} \frac{\frac{\beta_{n,\ell}}{\overline{z_n}(\ell-1)}}{(1 - \overline{z_n} z)^{\ell-1}} \tag{5.7.7}
$$

is an antiderivative of B on $\mathbb{C} \backslash \{1/\overline{z_1}, \ldots, 1/\overline{z_N}\}$. The constant α is arbitrary and we choose it so that $\mathfrak{B}(0) = 0$.

Since B has a zero of degree m at the origin, $\mathfrak{B}(0) = 0$, and $\mathfrak{B}' = B$, we conclude that $\mathfrak{B}$ has a zero of degree $m+1$ at the origin. Taking the common denominator in (5.7.7), we see that

$$
\mathfrak{B}(z) = \frac{z^{m+1} P(z)}{\prod_{n=1}^{N} (1 - \overline{z_n} z)^{m_n - 1}}, \tag{5.7.8}
$$

where P is a polynomial of degree at most $\sum_{n=1}^{N} (m_n - 1)$. On the other hand, (5.7.6) and (5.7.7) imply that

$$
\lim_{z \to \infty} \frac{(m+1)\mathfrak{B}(z)}{zB(z)} = 1. \tag{5.7.9}
$$

Define

$$
f(z) = \frac{(m+1)\mathfrak{B}(z)}{zB(z)}
$$

and use (5.7.2) and (5.7.8) to obtain

$$f(z) = \frac{(m+1)P(z)\prod_{n=1}^{N}(1-\overline{z_n}z)}{\prod_{n=1}^{N}(z-z_n)^{m_n}},$$

which reveals that f is analytic on $\mathbb{C}\backslash\mathbb{D}$. Since B has at least one pole in $\mathbb{C}\backslash\mathbb{D}^-$, f has at least one zero in $\mathbb{C}\backslash\mathbb{D}^-$. This enables us to produce a contradiction as follows. By (5.7.9) we know that

$$\lim_{z\to\infty} f(z) = 1,$$

and by (5.7.5),

$$\begin{aligned} |f(e^{i\theta})| &= \left|\frac{(m+1)\mathfrak{B}(e^{i\theta})}{e^{i\theta}B(e^{i\theta})}\right| \\ &= \left|(m+1)\mathfrak{B}(e^{i\theta})\right| \\ &\leqslant 1 \end{aligned}$$

for each $e^{i\theta} \in \mathbb{T}$. Since f is analytic on $\mathbb{C}\backslash\mathbb{D}$, $|f(\zeta)| \leqslant 1$ for $\zeta \in \mathbb{T}$, and

$$\lim_{z\to\infty} f(z) = 1,$$

the Maximum Modulus Principle ensures that $f \equiv 1$. This contradicts the assumption that f has zeros in $\mathbb{C}\backslash\mathbb{D}$. Thus, for some $n \in \{1, 2, \ldots, N\}$, we must have $\beta_{n,1} \neq 0$. □

5.8 Exercises

5.1 Verify (5.3.5) and (5.3.6).

5.2 Let

$$B(z) = z\prod_{k=1}^{n-1}\frac{z_k - z}{1-\overline{z}_k z},$$

in which $z_1, z_2, \ldots, z_{n-1} \in \mathbb{D}$ are distinct, let $w \in \mathbb{T}$, and let ζ_r be any of the n distinct solutions of $B(\zeta) = w$. Show that there is λ_r with $0 < \lambda_r < 1$ and a finite Blaschke product C of degree $n-1$ with $C(0) = 0$ such that

$$\frac{B(z)/z}{B(z)-w} = \frac{\lambda_r}{z-\zeta_r} + (1-\lambda_r)\frac{C(z)/z}{C(z)-w}.$$

Hint: Using the notation of Lemma 5.2.7, consider

$$R(z) = \frac{1}{1-\lambda_r}\left(\frac{B(z)}{B(z)-w} - \frac{\lambda_r z}{z-\zeta_r}\right),$$

and then define

$$C(z) = \frac{wR(z)}{R(z)-1}.$$

Note that R has $n-1$ simple poles on $\mathbb{T}$ and

$$\text{Re } R(z) = \frac{1}{2}, \qquad z \in \mathbb{T}.$$

5.3 In Theorem 5.7.1 the reader should be aware, as was pointed out by Heins [72] in this following construction, that there exist finite Blaschke products with at least two finite poles whose residue at some finite pole vanishes. Construct such an example as follows. Let $0 < a_1 < a_2 < 1$ and define

$$B(z) = \left(\frac{z-a_1}{1-a_1 z}\right)^2 \left(\frac{z-a_2}{1-a_2 z}\right)^2.$$

Fix a_1 and adjust the a_2 so that the residue at $1/a_1$ vanishes.

Chapter 6
Critical Points

In this chapter we consider the set of *critical points* $\{z : B'(z) = 0\}$ of a finite Blaschke product B. We first discuss their location, in terms of the zeros of B, and then we discuss the possibility of creating a finite Blaschke product with a desired set of critical points.

6.1 Location of the Critical Points

If B is a finite Blaschke product of degree n, then Theorem 3.3.2 and the quotient rule for derivatives ensure that $B' = P/Q$, in which P and Q are polynomials and $\deg P \leqslant 2n-2$. Lemma 3.4.3 implies that there are no zeros of B' on $\mathbb{T}$. They are either in $\mathbb{D}$ or in $\mathbb{D}_e$. In fact, we have the following symmetry result.

Lemma 6.1.1 *Let B be finite Blaschke product. Suppose that $w \in \mathbb{C}\backslash\{0\}$, $B(w) \neq 0$, and $B(w) \neq \infty$. Then $B'(w) = 0$ if and only if $B'(1/\overline{w}) = 0$.*

Proof For each $z \in \mathbb{C}\backslash\{0\}$, (3.1.6) tells us that

$$B(z)\overline{B(1/\overline{z})} = 1. \tag{6.1.2}$$

This implies

$$B(w) \neq 0 \qquad \Longleftrightarrow \qquad B(1/\overline{w}) \neq \infty. \tag{6.1.3}$$

Taking the derivative with respect to z of the expression in (6.1.2) reveals that

$$B'(z)\overline{B(1/\overline{z})} - \frac{1}{z^2}B(z)\overline{B'(1/\overline{z})} = 0,$$

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_6

and hence, using (6.1.3), we have

$$B'(w) = 0 \quad \Longleftrightarrow \quad B'(1/\overline{w}) = 0. \qquad \square$$

Be mindful of the hypotheses $B(w) \neq 0$ and $B(w) \neq \infty$ when applying this lemma. For example, if

$$B(z) = \left(\frac{z-\alpha}{1-\overline{\alpha}z}\right)^2, \quad \alpha \in \mathbb{D}\backslash\{0\},$$

then $B'(\alpha) = 0$ but $B'(1/\overline{\alpha}) = \infty$. This fact is crystallized in the following theorem.

Theorem 6.1.4 *Let B be a finite Blaschke product of degree n. Write*

$$B(z) = e^{i\beta} z^{j_0} \prod_{k=1}^{m} \left(\frac{z_k - z}{1-\overline{z_k}z}\right)^{j_k},$$

in which $\beta \in [0, 2\pi)$, $j_0 \geqslant 0$, $j_1, j_2, \ldots, j_m$ are positive integers with

$$j_0 + j_1 + \cdots + j_m = n,$$

and $z_1, z_2, \ldots, z_m$ are distinct points in $\mathbb{D}\backslash\{0\}$. Then B' has $n-1$ zeros in $\mathbb{D}$ (counting multiplicity). The number of zeros in $\mathbb{C}\backslash\mathbb{D}^-$ is m if $j_0 > 0$ and less than or equal to $m-1$ if $j_0 = 0$.

Proof We again remind the reader that there are no critical points of B on $\mathbb{T}$ (Lemma 3.4.3). First suppose that the zeros of B are distinct and that neither B nor B' have any zeros at the origin. By (3.4.6), $B'(z) = 0$ if and only if

$$\sum_{k=1}^{n} \frac{1-|z_k|^2}{(1-\overline{z_k}z)(z-z_k)} = 0.$$

Multiplying both sides of the preceding by

$$\prod_{k=1}^{n} (1-\overline{z_k}z)(z-z_k),$$

we obtain a polynomial equation of degree $2n-2$ whose zeros are not in

$$\{0, z_1, z_2, \ldots, z_n, 1/\overline{z_1}, 1/\overline{z_2} \ldots, 1/\overline{z_n}\}.$$

By Lemma 6.1.1, there are exactly $n-1$ zeros in $\mathbb{D}$ and $n-1$ zeros in $\mathbb{C}\backslash\mathbb{D}^-$.

In the general case, Theorem 4.4.1 permits us to approximate B by a family B_ϵ of finite Blaschke products of degree n with distinct zeros so that neither B_ϵ nor B'_ϵ have any zeros at the origin (the convergence is uniform on compact subsets of $\mathbb{D}$). It follows that B' has exactly $n-1$ zeros, counted according to multiplicity, in $\mathbb{D}$. However, it may have fewer zeros in $\mathbb{D}_e$.

We now consider the zeros of B' in $\mathbb{D}_e$. First assume that $j_0 \neq 0$. By direct verification,

$$B'(z) = z^{j_0-1} \frac{\prod_{k=1}^{m}(z-z_k)^{j_k-1}}{\prod_{k=1}^{m}(z-1/\overline{z_k})^{j_k+1}} P(z),$$

in which P is a polynomial of degree $2m$ with no zeros in $\{0, z_1, \ldots, z_m\}$. As a result, B' has $n+m-1$ zeros in $\mathbb{C}$. These are the zeros of B and of P, repeated according to multiplicity. Lemma 6.1.1 implies that the zeros of P are of the form

$$w_1,\ 1/\overline{w_1},\ w_2,\ 1/\overline{w_2}, \ldots,\ w_m,\ 1/\overline{w_m},$$

in which $w_1, w_2, \ldots, w_m \in \mathbb{D}\backslash\{0, z_1, z_2, \ldots, z_m\}$.

Now suppose that $j_0 = 0$ and write

$$B(z) = C \frac{\prod_{k=1}^{m}(z-z_k)^{j_k}}{\prod_{k=1}^{m}(z-1/\overline{z_k})^{j_k}} = C\left(1 + \frac{Q(z)}{\prod_{k=1}^{m}(z-1/\overline{z_k})^{j_k}}\right),$$

in which C is constant and Q is a polynomial of degree $n-1$. Thus,

$$B'(z) = \frac{\prod_{k=1}^{m}(z-z_k)^{j_k-1}}{\prod_{k=1}^{m}(z-1/\overline{z_k})^{j_k+1}} P(z),$$

in which P is a polynomial of degree at most $2m-2$ that has no zeros among $\{z_1, z_2, \ldots, z_m\}$. Consequently, B' has at most $n+m-2$ zeros in $\mathbb{C}$. These are the zeros of B and of P, repeated according to multiplicity. In this case, P might have zeros at the origin. For the rest of its zeros, Lemma 6.1.1 applies. Therefore, P can have, say, ℓ zeros at the origin and the rest are of the form

$$w_1,\ 1/\overline{w_1},\ w_2,\ 1/\overline{w_2},\ \ldots,\ w_{\ell'},\ 1/\overline{w_{\ell'}},$$

where $w_1, w_2, \ldots, w_{\ell'} \in \mathbb{D}\backslash\{0, z_1, z_2, \ldots, z_m\}$. Since $\ell' + 2\ell = \deg P \leqslant 2m-2$, we have $\ell' \leqslant m-1$. □

Recall that Corollary 5.1.4 (commonly known as the Gauss–Lucas theorem, although we have reserved that name for the more general Theorem 5.1.1) asserts that if P is a nonconstant polynomial, then the zeros of P' are contained in the convex hull of the zeros of P. An analogous result holds for the zeros of the derivative of a finite Blaschke product B: the zeros of B' are in the convex hull of $B^{-1}(\{0\}) \cup \{0\}$ [18]. A refinement of this result from [49] (see also [129, 137]) involves some hyperbolic geometry.

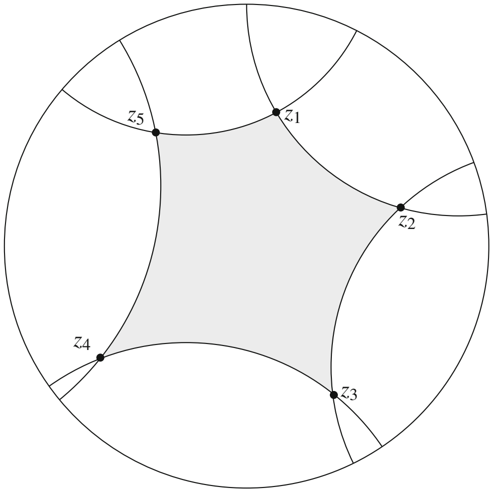


Fig. 6.1 The hyperbolic convex hull of the set $\{z_1, z_2, z_3, z_4, z_5\}$

Definition 6.1.5 We say that $A \subseteq \mathbb{D}$ is *hyperbolically convex* if

$$z_1, z_2 \in A \text{ and } t \in [0,1] \quad \Longrightarrow \quad \frac{z_1 - \frac{z_1 - z_2}{1 - \overline{z_1} z_2} t}{1 - \overline{z_1} \frac{z_1 - z_2}{1 - \overline{z_1} z_2} t} \in A.$$

The complicated quotient above is the parametric formula for the hyperbolic line segment between z_1 and z_2; see (2.3.10). The *hyperbolic convex hull* of $A \subseteq \mathbb{D}$ is the smallest hyperbolically convex set that contains A. It is the intersection of all hyperbolically convex sets that contain A. Figure 6.1 shows the hyperbolic convex hull of a set of five points in $\mathbb{D}$.

Theorem 6.1.6 *If B is a finite Blaschke product, then the zeros of B' in $\mathbb{D}$ belong to the hyperbolic convex hull of the zeros of B.*

Proof Let $\mathbb{D}_+ = \mathbb{D} \cap \{z : \operatorname{Im} z > 0\}$ and $\mathbb{D}_- = \mathbb{D} \cap \{z : \operatorname{Im} z < 0\}$ denote the upper and lower half disks, respectively. Suppose that the zeros of B all belong to $\mathbb{D}_+$. By (3.4.6), we have

$$\operatorname{Im}\left(\frac{B'(z)}{B(z)}\right) = \sum_{k=1}^{n} \operatorname{Im}\left(\frac{1 - |z_k|^2}{(1 - \overline{z_k} z)(z - z_k)}\right). \tag{6.1.7}$$

Let

$$\varphi(z) = \frac{1 - |a|^2}{(1 - \overline{a} z)(z - a)},$$

in which $a \in \mathbb{D}_+$ is fixed.

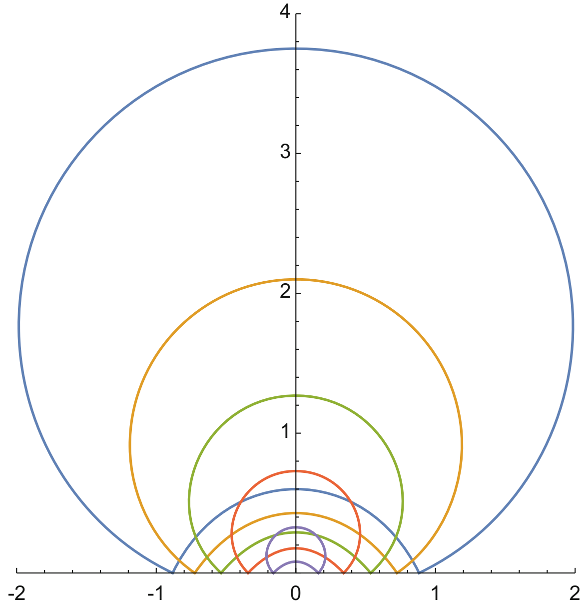


Fig. 6.2 The image of $\mathbb{T}_-$ and the interval $[-1, 1]$ under the mapping φ for $a = 0.25, 0.4, 0.55, 0.70, 0.85$

To study $\varphi(\mathbb{D}_-)$, we examine the image of the boundary

$$\{e^{i\theta} : -\pi \leqslant \theta \leqslant 0\} \cup [-1, 1]$$

of $\mathbb{D}_-$ under φ (see Fig. 6.2). On $\mathbb{T}_- = \{e^{i\theta} : -\pi \leqslant \theta \leqslant 0\}$ we have

$$\begin{aligned}\varphi(e^{i\theta}) &= \frac{1 - |a|^2}{(1 - \overline{a}e^{i\theta})(e^{i\theta} - a)}\\ &= \frac{1 - |a|^2}{|e^{i\theta} - a|^2} e^{-i\theta}\end{aligned}$$

and hence $\mathbb{T}_-$ is mapped onto a curve in $\mathbb{C}_+ \cup \mathbb{R}$. For $t \in [-1, 1]$,

$$\begin{aligned}\varphi(t) &= \frac{1 - |a|^2}{(1 - \overline{a}t)(t - a)}\\ &= \frac{1 - |a|^2}{|(1 - \overline{a}t)(t - a)|^2}(1 - at)(t - \overline{a}),\end{aligned}$$

and hence

$$\begin{aligned}\operatorname{Im}\varphi(t) &= \operatorname{Im}\frac{1 - |a|^2}{(1 - \overline{a}t)(t - a)}\\ &= \frac{1 - |a|^2}{|(1 - \overline{a}t)(t - a)|^2}(1 - t^2)\operatorname{Im} a.\end{aligned}$$

Therefore, $[-1, 1]$ is also mapped to a curve in $\mathbb{C}_+ \cup \mathbb{R}$ and hence the boundary of $\mathbb{D}_-$ is mapped to a simple closed curve in $\mathbb{C}_+ \cup \mathbb{R}$. Since φ is analytic on $(\mathbb{D}_-)^-$, we deduce that φ maps $\mathbb{D}_-$ into $\mathbb{C}_+$. Equivalently,

$$z \in \mathbb{D}_- \quad \Longrightarrow \quad \operatorname{Im} \varphi(z) > 0.$$

Since the zeros of B are in $\mathbb{C}_+$, the representation (6.1.7) implies

$$z \in \mathbb{D}_- \quad \Longrightarrow \quad \operatorname{Im} \left(\frac{B'(z)}{B(z)} \right) > 0.$$

Hence B' does not have any zeros in $\mathbb{D}_-$. By continuity, it follows that if all zeros of B are in $\mathbb{D}_+ \cup (-1, 1)$, then so are the zeros of B' (recall that we only consider the zeros inside $\mathbb{D}$).

Let $f = B \circ \tau_a$. By Lemma 3.6.1, f is also a finite Blaschke product with zeros $\tau_a(z_1), \tau_a(z_2), \dots, \tau_a(z_n)$ (Fig. 6.3). If we denote the zeros of B' in $\mathbb{D}$ by $w_1, w_2, \dots, w_{n-1}$, then the zeros of f' in $\mathbb{D}$ are

$$\tau_a(w_1),\ \tau_a(w_2), \dots, \tau_a(w_{n-1}).$$

If we choose a such that $\operatorname{Im} \tau_a(z_k) \geqslant 0$ for $1 \leqslant k \leqslant n$, then the preceding observation shows that

$$\operatorname{Im} \tau_a(w_k) \geqslant 0, \quad 1 \leqslant k \leqslant n-1.$$

This means that if the zeros of B are on one side of the hyperbolic line

$$\frac{a-z}{1-\bar{a}z} = t, \quad t \in [-1, 1],$$

then the zeros of B' are also on the same side. Similar comments apply if we replace τ_a by a rotation ρ_γ. The intersection of all such hyperbolic planes gives the hyperbolic convex hull of the zeros of B. □

Example 6.1.8 Let $a, b \in \mathbb{D}$ be unequal and let

$$B(z) = \left(\frac{a-z}{1-\bar{a}z} \right)^m \left(\frac{b-z}{1-\bar{b}z} \right)^n.$$

Then B' has $m+n-1$ zeros in $\mathbb{D}$. To be more specific, they are a (with multiplicity $m-1$), b (with multiplicity $n-1$), and c (with multiplicity one), which is the solution of the equation

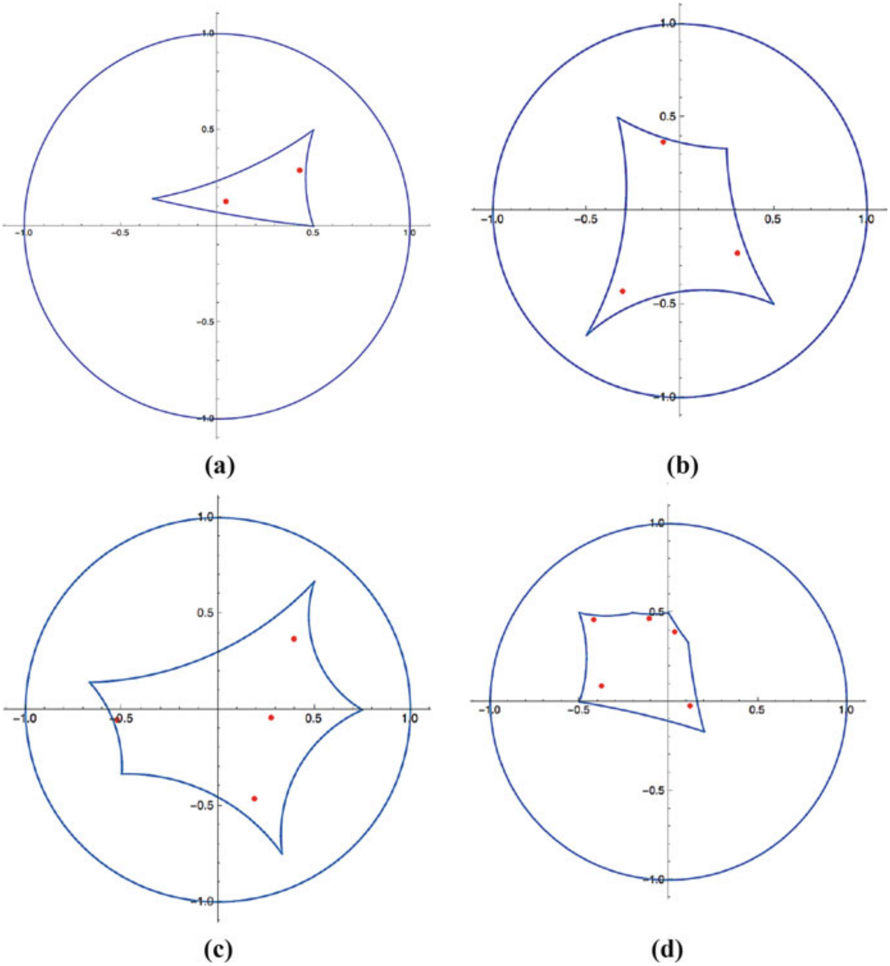


Fig. 6.3 (**a**)–(**d**) depict the hyperbolic convex hulls of the zeros of finite Blaschke products B of degrees 3, 4, 5, 6 (respectively) along with the zeros of B'. Observe how the zeros of B' lie within the hyperbolic convex hull of the zeros of B. We thank Tongzhou Wang and Raymone Cao for rendering these drawings

$$\frac{m(1-|a|^2)}{(1-\overline{a}z)^2}\left(\frac{a-z}{1-\overline{a}z}\right)^{m-1}\left(\frac{b-z}{1-\overline{b}z}\right)^{n} + \left(\frac{a-z}{1-\overline{a}z}\right)^{m}\left(\frac{b-z}{1-\overline{b}z}\right)^{n-1}\frac{n(1-|b|^2)}{(1-\overline{b}z)^2} = 0.$$

Since the above can be written as

$$\left(\frac{z-a}{1-a\overline{z}}\right)\Big/\left(\frac{z-b}{1-b\overline{z}}\right) = -\left(\frac{m(1-|a|^2)}{|1-\overline{a}z|^2}\right)\Big/\left(\frac{n(1-|b|^2)}{|1-\overline{b}z|^2}\right),$$

Corollary 2.3.11 ensures that a, b, and c lie on the same hyperbolic line. Moreover, as m and n vary independently over the positive integers, the point c traverses a dense subset of the hyperbolic line segment between a and b; see Exercise 6.3.

6.2 Controlling the Critical Points

Theorem 6.1.4 says that a Blaschke product of order $d+1$ has d critical points. The following theorem of Heins [74] shows that one has complete freedom to choose the location of these critical points.

Theorem 6.2.1 *Let $c_1, c_2, \ldots, c_d$ be d (not necessarily distinct) points in $\mathbb{D}$. Then there is a unique finite Blaschke product f of degree $d+1$ with $f(0) = 0$ and $f(1) = 1$ and having $c_1, c_2, \ldots, c_d$ as its critical points. Moreover, if g is any other finite Blaschke product of degree $d+1$ with critical points $c_1, c_2, \ldots, c_d$, then there is a $\tau \in \mathrm{Aut}(\mathbb{D})$ such that*

$$\tau \circ g = f.$$

Our proof follows Zakeri [139] and requires some preliminary ideas from point set and algebraic topology. Informally, the main idea is to show that the map

$$\{0, z_1, z_2, \ldots, z_d\} \mapsto \{c_1, c_2, \ldots, c_d\}$$

which takes the zeros $\{0, z_1, z_2, \ldots, z_d\}$ of a finite Blaschke product f of degree $d+1$ with $f(0) = 0$ and $f(1) = 1$ to the critical points $\{c_1, c_2, \ldots, c_d\}$ of a finite Blaschke product of degree $d+1$ is onto. We will do this by defining certain quotient topologies on the domain and range spaces of this map. To get started, let us first define some topological notions on the set of finite Blaschke products of degree d.

For sequences $\gamma_n \in \mathbb{T}$ and $a_n \in \mathbb{D}$, consider the sequence of disk automorphisms

$$\gamma_n \frac{z-a_n}{1-\overline{a_n}z}.$$

If $\gamma_n \to \zeta \in \mathbb{T}$ and $a_n \to a \in \mathbb{D}$, then one can see that the sequence of automorphisms above converges uniformly on compact subsets of $\mathbb{D}$ to the automorphism

$$\gamma \frac{z-a}{1-\overline{a}z}.$$

Therefore, if

$$B_n(z) = \gamma_n \prod_{j=1}^{d} \frac{z - a_{j,n}}{1 - \overline{a_{j,n}} z}$$

is a finite Blaschke product of degree d with $\gamma_n \to \zeta \in \mathbb{T}$ and $a_{j,n} \to a_j \in \mathbb{D}$ as $n \to \infty$, then B_n converges uniformly on compact subsets of $\mathbb{D}$ to the finite Blaschke product

$$B(z) = \gamma \prod_{j=1}^{d} \frac{z - a_j}{1 - \overline{a_j} z}.$$

Under the circumstances above, the degree of B_n remains invariant. Indeed, the limiting finite Blaschke product has the same degree d as each of the B_n.

Now let us consider the case when some of the zeros $a_{j,n}$ tend to a point on $\mathbb{T}$. Here the situation changes and the interplay between the zeros $a_{j,n}$ and the unimodular constant γ_n becomes important. This leads us to consider following normalized disk automorphisms. For each $a \in \mathbb{D}$, define

$$\beta(a, z) := \frac{1 - \overline{a}}{1 - a} \cdot \frac{z - a}{1 - \overline{a} z}. \tag{6.2.2}$$

Observe that $\beta(a, z)$ is the unique element of $\operatorname{Aut}(\mathbb{D})$ for which

$$\beta(a, 1) = 1 \quad \text{and} \quad \beta(a, a) = 0.$$

The following proposition focuses on what happens to $\beta(a, z)$ when the parameter $a \in \mathbb{D}$ approaches a point of $\mathbb{T}$.

Proposition 6.2.3 *Suppose that a_n is a sequence in $\mathbb{D}$ and let $\beta(a_n, z)$ be the corresponding sequence of disk automorphisms defined by* (6.2.2).

(a) If $a_n \to a \in \mathbb{T} \backslash \{1\}$, then $\beta(a_n, z) \to 1$ uniformly on compact subsets of $\mathbb{D}$.
(b) If $a_n \to 1$, then for each accumulation point $\alpha \in \mathbb{T}$ of the sequence

$$\frac{1 - \overline{a_n}}{1 - a_n},$$

there is a subsequence of the $\beta(a_n, z)$ that converges to the constant function $-\alpha$ uniformly on compact subsets of $\mathbb{D}$. In particular, if

$$\frac{1 - \overline{a_n}}{1 - a_n} \to \alpha,$$

then $\beta(a_n, z)$ converges to $-\alpha$ uniformly on compact subsets of $\mathbb{D}$.

Proof

(a) Fix a compact set $K \subseteq \mathbb{D}$ and $z \in K$. Since a_n is bounded away from 1 we have

$$|\beta(a_n, z) - 1| = \frac{|z-1|}{|1-a_n||1-\overline{a_n}z|}(1-|a_n|^2) \leqslant C_K(1-|a_n|^2),$$

which goes to zero as $n \to \infty$.

(b) Without loss of generality assume that

$$\lim_{n\to\infty} \frac{1-\overline{a_n}}{1-a_n} \to \alpha.$$

Then for each $z \in K$,

$$\begin{aligned}|\beta(a_n, z) + \alpha| &= \left| \frac{1-\overline{a_n}}{1-a_n}\left(\frac{z-a_n}{1-\overline{a_n}z} + 1\right) + \left(\alpha - \frac{1-\overline{a_n}}{1-a_n}\right)\right| \\ &\leqslant \frac{|z|+1}{|1-\overline{a_n}z|}|1-a_n| + \left|\alpha - \frac{1-\overline{a_n}}{1-a_n}\right| \\ &\leqslant C_K|1-a_n| + \left|\alpha - \frac{1-\overline{a_n}}{1-a_n}\right|,\end{aligned}$$

which goes to zero as $n \to \infty$. □

Let $\mathscr{B}_d$ denote the family of all finite Blaschke products of the form

$$B(z) = z\prod_{j=1}^{d} \beta(a_n, z). \tag{6.2.4}$$

Each element of $\mathscr{B}_d$ is of degree $d+1$ and is normalized so that

$$B(0) = 0 \quad \text{and} \quad B(1) = 1.$$

Proposition 6.2.3 yields the following corollary.

Corollary 6.2.5 *For a sequence $B_n \in \mathscr{B}_d$, either B_n converges uniformly on compact subsets of $\mathbb{D}$ to some $B \in \mathscr{B}_d$ or each subsequence of B_n has a subsequence that converges uniformly on compact subsets of $\mathbb{D}$ to γB for some $\gamma \in \mathbb{T}$ and $B \in \mathscr{B}_{d'}$ with $0 \leqslant d' < d$.*

6.3 The Topological Space Σ_d

For an equivalence relation $\sim$ on a topological space X, let

$$[x] = \{x' \in X : x' \sim x\}$$

denote the equivalence class of x and

$$X/\sim\ = \{[x] : x \in X\}$$

the set of equivalence classes. If

$$\pi : X \to X/\sim, \quad \pi(x) = [x] \tag{6.3.1}$$

is the canonical projection map, then

$$\{V \subseteq X/\sim : \pi^{-1}(V) \text{ is open in } X\}$$

is a collection of open sets that defines the *quotient topology* on $X/\sim$. The resulting topological space $X/\sim$ is a *quotient space*.

Given $A \subseteq X$, its *saturation* is the set

$$\pi^{-1}(\pi(A)).$$

In other words, the saturation of A is the collection of all elements of X that are related to some element of A via $\sim$. As a consequence of the definitions, we see that π is an open mapping if and only if the saturation of each open subset of X is open.

We now apply the general construction above to the polydisk

$$\mathbb{D}^d = \{(z_1, z_2, \ldots, z_d) : z_j \in \mathbb{D}\},$$

endowed with the product topology. That is, given

$$(z_1, z_2, \ldots, z_d) \in \mathbb{D}^d,$$

a local basis for the Cartesian (product) topology is the collection of sets

$$V_1 \times \cdots \times V_d,$$

where $V_j \subseteq \mathbb{D}$ is an open neighborhood of z_j. Let S_d be the symmetric group on the set $\{1, 2, \ldots, d\}$, that is, the set of bijective mappings of $\{1, 2, \ldots, d\}$ to itself. We define an equivalence relation $\sim$ on $\mathbb{D}^d$ by setting

$$(a_1, a_2, \ldots, a_d) \sim (b_1, b_2, \ldots, b_d)$$

if there is a permutation $\sigma \in S_d$ such that

$$b_j = a_{\sigma(j)}, \quad 1 \leqslant j \leqslant d.$$

Then the quotient space $\mathbb{D}/\sim$, denoted by Σ_d, can be thought of as the set of *unordered* d-tuples (a multiset) $\langle a_1, a_2, \ldots, a_d\rangle$, with $a_j \in \mathbb{D}$.

We use the bracket notation $\langle a_1, a_2, \ldots, a_d\rangle$ rather than the standard set notation $\{a_1, a_2, \ldots, a_d\}$, since we are not concerned with the order of how the a_js are listed in the set and we are allowing repetitions of the a_js. For example, we permit $\langle 0, 0, \frac{1}{2}\rangle$ to be an element of Σ_3. If we use the conventional set notation, the set $\{0, 0, \frac{1}{2}\}$ is the same as the set $\{0, \frac{1}{2}\}$. In our application of this, the multiset $\langle a_1, a_2, \ldots, a_d\rangle$ denotes the zeros of a finite Blaschke product so the order in which we list the zeros does not matter but the repetitions do matter.

The quotient space Σ_d plays a crucial role in our study of the critical points of finite Blaschke products. In what follows, π denotes the canonical projection of $\mathbb{D}^d$ onto Σ_d from (6.3.1).

A local base at the point $a = \langle a_1, a_2, \ldots, a_d\rangle \in \Sigma_d$ is obtained as follows. Given $\varepsilon > 0$, consider the saturated open set

$$V_\varepsilon := \bigcup_{\sigma \in S_d} \big(D(a_{\sigma(1)}, \varepsilon) \times D(a_{\sigma(2)}, \varepsilon) \times \cdots \times D(a_{\sigma(d)}, \varepsilon)\big) \subseteq \mathbb{D}^d,$$

where $D(a_j, \epsilon)$ is the open disk of radius ϵ centered at a_j and the radius ε is taken small enough so that all of the disks $D(a_j, \varepsilon)$ remain in $\mathbb{D}$. Then $\pi(V_\varepsilon)$ is an open neighborhood of $a \in \Sigma_d$ which, for shorthand, we denote by $\mathfrak{D}(a, \varepsilon)$. In other words, $\mathfrak{D}(a, \varepsilon)$ is the set of multisets $b = \langle b_1, b_2, \ldots, b_d\rangle \in \Sigma_d$ for which there is a permutation $\sigma \in S_d$ such that

$$|a_1 - b_{\sigma(1)}| < \varepsilon, \quad |a_2 - b_{\sigma(2)}| < \varepsilon, \ldots, \quad |a_d - b_{\sigma(d)}| < \varepsilon.$$

Example 6.3.2 If $a = \langle 0, 0, \frac{1}{2}\rangle \in \Sigma_3$, then

$$\begin{aligned}\mathfrak{D}(a, \varepsilon) = \Big\{ b = \langle b_1, b_2, b_3\rangle \in \Sigma_3 : {}& |b_1| < \varepsilon,\ |b_2| < \varepsilon,\ |b_3 - 1/2| < \varepsilon, \\ & \text{or } |b_1| < \varepsilon,\ |b_2 - 1/2| < \varepsilon,\ |b_3| < \varepsilon, \\ & \text{or } |b_1 - 1/2| < \varepsilon,\ |b_2| < \varepsilon,\ |b_3| < \varepsilon \Big\}.\end{aligned}$$

The inverse image of $\mathfrak{D}(a, \varepsilon)$ under π is

$$\begin{aligned}V_\varepsilon = {}& \Big\{ \{|z| < \varepsilon\} \times \{|z| < \varepsilon\} \times \{|z - 1/2| < \varepsilon\} \Big\} \\ & \cup \Big\{ \{|z| < \varepsilon\} \times \{|z - 1/2| < \varepsilon\} \times \{|z| < \varepsilon\} \Big\} \\ & \cup \Big\{ \{|z - 1/2| < \varepsilon\} \times \{|z| < \varepsilon\} \times \{|z| < \varepsilon\} \Big\}.\end{aligned}$$

This observation enables us to characterize convergent sequences in Σ_d. We see that $a_n = \langle a_{n,1}, a_{n,2}, \ldots, a_{n,d} \rangle$ converges to $a = \langle a_1, a_2, \ldots, a_d \rangle$ in Σ_d if and only if for each $\varepsilon > 0$ there is an $N = N(\varepsilon)$ such that for each $n \geqslant N$ there is a permutation $\sigma \in S_d$ (which may depend on n) such that

$$|a_{n,1} - a_{\sigma(1)}| < \varepsilon, \quad |a_{n,2} - a_{\sigma(2)}| < \varepsilon, \ldots, \ |a_{n,d} - a_{\sigma(d)}| < \varepsilon.$$

Proposition 6.3.3 *π is an open mapping, meaning that if U is an open subset of $\mathbb{D}^d$ in the product topology, then $\pi(U)$ is an open subset of Σ_d in the quotient topology.*

Proof For each fixed $\sigma \in S_d$, the mapping

$$F_\sigma : \mathbb{D}^d \to \mathbb{D}^d, \quad F_\sigma(a_1, a_2, \ldots, a_d) = (a_{\sigma(1)}, a_{\sigma(2)}, \ldots, a_{\sigma(d)}),$$

is a homeomorphism of $\mathbb{D}^d$ (endowed with its natural product topology). Since for each subset $A \subseteq \mathbb{D}^d$,

$$\pi^{-1}(\pi(A)) = \bigcup_{\sigma \in S_d} F_\sigma(A),$$

the saturation of each open set is open. Hence, π is an open mapping. □

Let us discuss an equivalent interpretation of Σ_d. For $a = \langle a_1, a_2, \ldots, a_d \rangle \in \Sigma_d$, define the corresponding finite Blaschke product $B(a, z) \in \mathscr{B}_d$ by

$$B(a, z) = z \prod_{j=1}^{d} \beta(a_j, z). \tag{6.3.4}$$

As discussed in the previous section, if $a_n = \{a_{n,1}, a_{n,2}, \ldots, a_{n,d}\}$ is a sequence in Σ_d, then $a_n \to a = \{a_1, a_2, \ldots, a_d\}$ in the topology of Σ_d if and only if for each $\varepsilon > 0$ there is an $N = N(\varepsilon)$ such that for each $n \geqslant N$ there is a permutation $\sigma \in S_d$ (which may depend on n) such that

$$|a_{n,1} - a_{\sigma(1)}| < \varepsilon, \quad |a_{n,2} - a_{\sigma(2)}| < \varepsilon, \ldots, \ |a_{n,d} - a_{\sigma(d)}| < \varepsilon.$$

Therefore, $a_n \to a$ in Σ_d implies that $B(a_n, z) \to B(a, z)$ uniformly on compact subsets of $\mathbb{D}$. Conversely, if the latter holds, then by Hurwitz's theorem, the zeros of $B(a_n, z)$ must be close to the zeros of $B(a, z)$. More precisely, for each $\varepsilon > 0$ small enough, there is an $N = N(\varepsilon)$ such that for each $n \geqslant N$

$$|a_{n,1} - a_{\sigma(1)}| < \varepsilon, \quad |a_{n,2} - a_{\sigma(2)}| < \varepsilon, \ldots, \ |a_{n,d} - a_{\sigma(d)}| < \varepsilon,$$

for a suitable permutation $\sigma \in S_d$. Therefore, $a_n \to a$ in the topology of Σ_d. In short, convergence of a sequence in Σ_d is equivalent to uniform convergence on compact subsets of $\mathbb{D}$ of the corresponding sequence (via (6.3.4)) of finite Blaschke products in $\mathscr{B}_d$.

The set $\mathscr{B}_d$ is endowed with the topology of uniform convergence on compact subsets of $\mathbb{D}$. Using the metric

$$d(f, g) = \sum_{n=2}^{\infty} \frac{1}{2^n} \sup \left\{ |f(z) - g(z)| : |z| \leqslant 1 - \frac{1}{n} \right\},$$

we see that $\mathscr{B}_d$, and hence Σ_d, via the identification (6.3.4), is metrizable [26, p. 143]. This observation is useful when discussing the critical points of a finite Blaschke product.

To get started, we define

$$\Phi : \Sigma_d \to \Sigma_d, \quad \Phi(a) = c, \tag{6.3.5}$$

where c is the unordered list of critical points of the finite Blaschke product $B(a, z)$. Note that $B(a, z)$ is a finite Blaschke product of order $d + 1$ and thus, counting multiplicities, it has precisely d critical points in $\mathbb{D}$ (Theorem 6.1.4). A key part of proving Theorem 6.2.1 (any set of d points in $\mathbb{D}$ can be the set of critical points of a finite Blaschke product of degree $d + 1$) is showing that Φ is onto.

Lemma 6.3.6 *Φ is continuous.*

Proof Let $a_n \to a$ in the topology of Σ_d. Thus, by the discussion above, $B(a_n, z) \to B(a, z)$ uniformly on compact subsets of $\mathbb{D}$. Therefore, by the Cauchy integral formula for the derivative,

$$\frac{d}{dz} B(a_n, z) \to \frac{d}{dz} B(a, z)$$

uniformly on compact subsets of $\mathbb{D}$. Applying Hurwitz' theorem [26, p. 152] to the critical points of $B(a_n, z)$, we conclude that $c_n \to c$ in the topology of Σ_d. □

As mentioned before, our ultimate goal is to show that Φ is a homeomorphism. We need another concept to do this.

Definition 6.3.7 For two metric spaces X and Y, a continuous mapping $f : X \to Y$ is *proper* if for each compact set $K \subseteq Y$, the inverse image $f^{-1}(K)$ is compact in X.

For more on proper mappings, see [98]. A sequential characterization of proper mappings is as follows.

Definition 6.3.8 A sequence $x_n \in X$ *escapes to infinity* if for each compact subset $K \subseteq X$, the set $\{n : x_n \in K\}$ is finite.

One can see that if x_n escapes to infinity, then so does every subsequence of x_n.

Proposition 6.3.9 *For $f : X \to Y$, a continuous map between two metric spaces X and Y, the following are equivalent.*

(a) f is proper.

(b) If $x_n \in X$ is any sequence that escapes to infinity, then $f(x_n)$ escapes to infinity in Y.

Proof $(a) \implies (b)$ Suppose x_n is a sequence in X that escapes to infinity. If $f(x_n)$ does not escape to infinity in Y, then there is a compact set $K \subseteq Y$ such that the cardinality of $\{n : f(x_n) \in K\}$ is infinite. Hence the cardinality of $\{n : x_n \in f^{-1}(K)\}$ is infinite. Since f is proper, $f^{-1}(K)$ is compact in X and this is a contradiction to the definition of escapes to infinity. Therefore, $f(x_n)$ escapes to infinity in Y.

$(b) \implies (a)$ Suppose K is a compact subset of Y. Since X and Y are metric spaces, it suffices to show that $f^{-1}(K)$ is sequentially compact (that is, if $x_n \in f^{-1}(K)$, then there is a subsequence that converges in $f^{-1}(K)$). To this end, let x_n be a sequence in $f^{-1}(K)$. Since $f(x_n) \in K$, it does not escape to infinity in Y. By assumption, x_n also does not escape to infinity in X. Hence, there is a compact set $L \subseteq X$ such that the cardinality of the set $\{n : x_n \in L\}$ is infinite. The sequential compactness of L provides a subsequence of x_n that converges in L. By continuity, $f^{-1}(K)$ is closed in X. Thus, x_n has a subsequence that converges to a point in $f^{-1}(K)$. □

Proposition 6.3.10 *Let X and Y be metric spaces and let $f : X \to Y$ be a proper continuous map. Then $f(X)$ is a closed subset of Y.*

Proof Suppose toward a contradiction that $y_0 \in f(X)^- \backslash f(X)$. Then there is a sequence $y_n \in f(X)$ such that $y_n \to y_0$. Therefore, the set

$$E = \{y_0\} \cup \{y_n : n \geqslant 1\}$$

is compact in Y. Since f is proper, $f^{-1}(E)$ compact in X. However, $y_0 \notin f(X)$ and hence

$$f^{-1}(E) = \{f^{-1}(\{y_n\}) : n \geqslant 1\}.$$

Take any $x_n \in f^{-1}(\{y_n\})$. This sequence, being in a compact subset of X, has a convergent subsequence. Without loss of generality, we may assume that x_n converges to $x_0 \in X$. By continuity, $y_0 = f(x_0) \in f(X)$, which is a contradiction. □

By Lemma 6.3.6, the mapping Φ from (6.3.5) is continuous. Using Proposition 6.3.9, we can say more.

Lemma 6.3.11 *$\Phi : \Sigma_d \to \Sigma_d$ is proper.*

Proof We will apply the criterion from Proposition 6.3.9. Suppose that $a_n = \langle a_{n,1}, a_{n,2}, \ldots, a_{n,d}\rangle$ is a sequence in Σ_d that escapes to infinity, but for which $\Phi(a_n)$ does not escape to infinity in Σ_d. By passing to a subsequence if necessary, we can assume that $\Phi(a_n)$ is confined to a compact subset of Σ_d. Since a_n escapes to infinity, by passing to a further subsequence and relabeling, we may assume that a_n converges to $a = \langle w_1, w_2, \ldots, w_d \rangle$, where $|w_1| = 1$ and $|w_j| \leqslant 1$ for the other indices j.

By Proposition 6.2.3, the corresponding sequence of normalized finite Blaschke products $B(a_n, z)$ has a subsequence that converges uniformly on compact subsets of $\mathbb{D}$ to a finite Blaschke product B of degree $d' + 1$ with $0 \leqslant d' < d$. In fact, the zeros of B are precisely at the origin and those w_j with $|w_j| < 1$. Therefore, B has d' critical points in $\mathbb{D}$ (Theorem 6.1.4).

However, uniform convergence on compact subsets of $\mathbb{D}$ and the fact that $\Phi(a_n)$ is confined to a compact subset of Σ_d implies that B has at least d critical points in $\mathbb{D}$. This contradicts the fact that B has $d' < d$ critical points. □

6.4 The Distance-Ratio Function

Let f belong to the Schur class $\mathscr{S}$ and endow $f(\mathbb{D})$ with the Poincaré metric

$$\frac{2|dw|}{1 - |w|^2}$$

from (2.4.1). Then its pullback under f is $\sigma_f(z)|dz|$, where

$$\sigma_f(z) = \frac{2|f'(z)|}{1 - |f(z)|^2}.$$

The metric $\sigma_f(z)|dz|$ has constant curvature -1 on $\mathbb{D}$, except at the critical points of f. This follows from the identity

$$\triangle \log \sigma_f(z) = \sigma_f^2(z). \tag{6.4.1}$$

From the Schwarz–Pick Lemma, f is contractive in the Poincaré metric, that is,

$$\frac{|f'(z)|}{1 - |f(z)|^2} \leqslant \frac{1}{1 - |z|^2}.$$

This leads us to consider the *distance-ratio function*

$$R_f(z) := \frac{1 - |z|^2}{1 - |f(z)|^2}|f'(z)| \tag{6.4.2}$$

which compares the pullback metric with the Poincaré metric. We gather some elementary properties of R_f below. Recall that a function $R : \mathbb{D} \to \mathbb{C}$ is *real analytic* if it can be represented as

$$R(z) = \sum_{m,n \geqslant 0} a_{m,n} z^m \overline{z^n},$$

where the partial sums of the series converges absolutely and uniformly on compact subsets of $\mathbb{D}$.

Lemma 6.4.3 *Let $f, g \in \mathscr{S}$.*

(a) $R_f \leqslant 1$*, with equality if and only if* $f \in \mathrm{Aut}(\mathbb{D})$.
(b) $R_{f\circ g} = (R_f \circ g)\, R_g$*. In particular, for every* $\tau \in \mathbb{D}$,

$$R_{\tau\circ f} = R_f \quad \text{and} \quad R_{f\circ\tau} = R_f \circ \tau.$$

(c) R_f *is a nonnegative function on* $\mathbb{D}$ *with zeros at the critical points of* f*. It is real analytic on* $\mathbb{D}$*, except at its zeros. Moreover, at each zero* c *of* R_f*, we have*

$$R_f(z) = |z-c|^m \widetilde{R}(z),$$

where $m \geqslant 1$ *is the order of the zero of* f' *at* c *and* $\widetilde{R}$ *is a positive real-analytic function in a neighborhood of* c.

Proof

(a) This follows from the Schwarz–Pick theorem (Theorem 1.4.1).
(b) This is a consequence of the chain rule. Indeed, by (6.4.2),

$$\begin{aligned} R_{f\circ g} &= \frac{1-|z|^2}{1-|(f\circ g)(z)|^2}|(f\circ g)'(z)| \\ &= \frac{1-|z|^2}{1-|(f(g(z))|^2}|f'(g(z))|\,|g'(z)| \\ &= \frac{1-|g(z)|^2}{1-|(f(g(z))|^2}|f'(g(z))|\frac{1-|z|^2}{1-|g(z)|^2}|g'(z)| \\ &= R_f(g(z))\, R_g(z) \\ &= (R_f\circ g)(z)\, R_g(z). \end{aligned}$$

Since $R_\tau = 1$ for any $\tau \in \mathrm{Aut}(\mathbb{D})$, the other identities follow.
(c) Suppose that f' has degree m at c. Then the expansion

$$f(z) = f(c) + f^{(m+1)}(c)(z-c)^{m+1} + \cdots$$

holds in a neighborhood of c and $f^{(m+1)}(c) \neq 0$. Thus,

$$f'(z) = (z-c)^m g(z),$$

where g is analytic and $g(c) \neq 0$. Plugging this into (6.4.2) yields the result. □

This next detail follows from Theorem 5.2.5. We will discuss an extended version of it with Theorem 6.5.2 below.

Lemma 6.4.4 *Let f be a finite Blaschke product. Then R_f has a continuous extension to $\mathbb{D}^-$ and*

$$\lim_{|z|\to 1} R_f(z) = 1.$$

Recall that if a twice continuously differentiable function F on a domain in the complex plane has a local maximum at a point z_0, then

$$\triangle F(z_0) \leqslant 0. \tag{6.4.5}$$

This fact from calculus is the main ingredient needed to show that the map Φ from (6.3) is injective.

Lemma 6.4.6 *$\Phi : \Sigma_d \to \Sigma_d$ is injective.*

Proof Let $a, b \in \Sigma_d$ be such that $\Phi(a) = \Phi(b)$. Let f and g be the corresponding finite Blaschke products in $\mathscr{B}_d$ from (6.2.4). Our assumption means that f and g have the same critical points. Consider the function

$$h(z) = \frac{R_f(z)}{R_g(z)}, \quad z \in \mathbb{D}.$$

Observe that off the critical points of f and g, the function h is real analytic. Moreover, any singularity of h will arise from a zero of R_g which, by Lemma 6.4.3(c), cancels out with a zero of R_f. Hence, h is a real-analytic function on $\mathbb{D}$. Lemma 6.4.4 tells us that h has a continuous extension to $\mathbb{D}^-$ and

$$\lim_{|z|\to 1} h(z) = 1.$$

Let us show that $h \leqslant 1$ on $\mathbb{D}$. Suppose to the contrary that h has a maximum at $z_0 \in \mathbb{D}$ with $h(z_0) > 1$. By (6.4.5),

$$\triangle \log h(z_0) \leqslant 0. \tag{6.4.7}$$

However, $h = R_f/R_g = \sigma_f/\sigma_g$ and thus, by (6.4.1),

$$\begin{aligned} \triangle \log h &= \triangle \log \sigma_f - \triangle \log \sigma_g \\ &= \sigma_f^2 - \sigma_g^2. \end{aligned}$$

The identity above holds off the critical points of f and g. By continuity, it holds everywhere. We have $h(z_0) > 1$, which can be rewritten as $\sigma_f(z_0) > \sigma_g(z_0)$. On the other hand, (6.4.7) says that $\sigma_f(z_0) \leqslant \sigma_g(z_0)$. Therefore, $\sigma_f(z_0) = \sigma_g(z_0)$, or equivalently, $h(z_0) = 1$, a contradiction. A similar argument shows that $1/h \geqslant 1$ and so $h \equiv 1$. From here we get

$$\frac{|f'(z)|}{1-|f(z)|^2} = \frac{|g'(z)|}{1-|g(z)|^2}, \quad z \in \mathbb{D}.$$

A theorem of Liouville (see [90, Thm. C]) says the identity above implies that $f \equiv g$ and so Φ is injective. □

Theorem 6.4.8 *$\Phi : \Sigma_d \to \Sigma_d$ is a homeomorphism.*

Proof By Lemmas 6.3.6, 6.3.11, and 6.4.6, Φ is continuous, proper, and injective. We also know from Proposition 6.3.10 that Φ has closed range. Brouwer's Invariance of Domain Theorem [12] says that $\Phi(\Sigma_d)$ is an open subset of Σ_d. Since Σ_d is connected and $\Phi(\Sigma_d)$ is both open and closed (and nonempty), we see that $\Phi(\Sigma_d) = \Sigma_d$, that is, $\Phi : \Sigma_d \to \Sigma_d$ is a homeomorphism. □

With all the heavy lifting complete, here is the proof of Theorem 6.2.1. The existence and uniqueness of f is precisely the bijectivity of Φ, which was proved in Theorem 6.4.8. For the second part, suppose g is a finite Blaschke product of degree $d+1$ with the same critical points as f. Since critical points do not change upon post-composing with a disk automorphism, we may choose $\tau \in \operatorname{Aut}(\mathbb{D})$ such that

$$(\tau \circ g)(0) = 0 \quad \text{and} \quad (\tau \circ g)(1) = 1.$$

Therefore, f and $\tau \circ g \in \mathscr{B}_d$ and, moreover, they have the same critical points. But, since the mapping Φ is injective we see that $f = \tau \circ g$ and the result follows.

6.5 A Characterization of Heins

We saw in Lemma 6.4.4 that if f is a finite Blaschke product, then the distance-ratio function

$$R_f(z) = \frac{(1-|z|^2)|f'(z)|}{1-|f(z)|^2}$$

satisfies

$$\lim_{|z| \to 1^-} R_f(z) = 1.$$

A theorem of Heins [74] (see Theorem 6.5.2 below) says that this condition characterizes the finite Blaschke products amongst the Schur class $\mathscr{S}$ functions. To state this theorem we need the following definition.

Definition 6.5.1 An analytic function $f : \mathbb{D} \to \mathbb{C}$ *has the asymptotic value a at the point* $\zeta \in \mathbb{T}$ if there is a curve Γ inside $\mathbb{D}$ which terminates at ζ such that

$$\lim_{\substack{z \to \zeta \\ z \in \Gamma}} f(z) = a.$$

For example, consider the function

$$f(z) = \exp\Big(-\frac{1+z}{1-z}\Big)$$

and notice that f belongs to $\mathscr{S}$. Furthermore,

$$\frac{1+\zeta}{1-\zeta} \in i\mathbb{R}, \quad \zeta \in \mathbb{T}\backslash\{1\}$$

and so for each $\zeta \in \mathbb{T}\backslash\{1\}$, f has an asymptotic value $f(\zeta) \in \mathbb{T}$ at each $\zeta \in \mathbb{T}\backslash\{1\}$. Since

$$\lim_{r \to 1^-} f(r) = 0,$$

we see that f has an asymptotic value of 0 at 1.

A theorem of Lindelöf [25] says that if $f \in H^\infty$ and f has the asymptotic value a at the point $\zeta \in \mathbb{T}$, then the nontangential limit (see (1.6.6)) of f at ζ is equal to a.

Theorem 6.5.2 (Heins [74]) *For $f \in \mathscr{S}$, the following are equivalent.*

(a) f is a nonconstant finite Blaschke product.
(b) For each sequence a_k in $\mathbb{D}$ with $a_k \to \gamma$ for some $\gamma \in \mathbb{T}$, the functions

$$\tau_{f(a_k)} \circ f \circ \tau_{a_k}$$

converge uniformly on compact subsets of $\mathbb{D}$ to a rotation.
(c) $\lim_{|z| \to 1} R_f(z) = 1$.
(d) f has no asymptotic values in $\mathbb{D}$ and has a finite set of critical points.

Moreover, if any of the conditions above hold, then the rotation promised in (b) is ρ_λ, where $\lambda = f'(\gamma)/|f'(\gamma)|$.

Proof $(a) \implies (b)$ If f is a finite Blaschke product of degree $n \geqslant 1$, then for each fixed $k \in \mathbb{N}$, the function

$$f_k = \tau_{f(a_k)} \circ f \circ \tau_{a_k} \tag{6.5.3}$$

is also a finite Blaschke product of degree n (Lemma 3.6.1). Let

$$z_{k,1}, z_{k,2}, \ldots, z_{k,n}$$

denote the n zeros of f_k and observe that these zeros are the solutions to the equation

$$f(\tau_{a_k}(z)) = f(a_k).$$

Now let $w_{k,j} = \tau_{a_k}(z_{k,j})$ and note that

$$w_{k,1}, w_{k,2}, \ldots, w_{k,n}$$

are the solutions to $f(w) = f(a_k)$. Number these so that $w_{k,1} = a_k$ for $k \in \mathbb{N}$.

Note that for each fixed $j = 1, 2, \ldots, n$,

$$\lim_{k\to\infty} |w_{k,j}| = 1,$$

and

$$|w_{k,i} - w_{k,j}| \geqslant \delta_k, \quad 1 \leqslant i < j \leqslant n,$$

for some constant δ_k. In fact, for any finite Blaschke product B whose zeros are $\xi_1, \xi_2, \ldots, \xi_n$ and any M with

$$\max\{|\xi_1|, |\xi_2|, \ldots, |\xi_n|\} < M < 1,$$

there is a constant $\delta = \delta(M, B) > 0$ such that, for any two distinct points z, w in the annulus $\{z : M \leqslant |z| \leqslant 1/M\}$, we have

$$B(z) = B(w) \implies |z - w| \geqslant \delta.$$

This uniform separation occurs because the annulus is free from the critical points of B (Lemma 3.4.3). More precisely, as $k \to \infty$, the $w_{k,j}$ tend to the n distinct (Theorem 3.4.10) solutions to

$$f(w) = f(\gamma).$$

By the argument used to prove Proposition 6.2.3, τ_{a_k} converges uniformly on compact subsets of $\mathbb{D}$ to γ. Since $\tau_{a_k}(z_{k,j}) = w_{k,j}$ for $j > 1$ and $w_{k,j}$ does not tend to γ as $k \to \infty$, we conclude that

$$\lim_{k\to\infty} |z_{k,j}| = \lim_{k\to\infty} |\tau_{a_k}^{-1}(w_{k,j})| = 1, \quad 2 \leqslant j \leqslant n. \tag{6.5.4}$$

Note that f_k is a finite Blaschke product with zeros at the origin and at $z_{k,j}$ for $2 \leqslant j \leqslant k$. Thus,

$$f_k(z) = \eta_k z \prod_{j=2}^{n} \frac{z_{k,j} - z}{1 - \overline{z_{k,j}} z},$$

where $\eta_k \in \mathbb{T}$. This formula yields

$$f_k'(0) = \eta_k \prod_{j=2}^{n} z_{k,j}$$

and hence

$$\frac{f_k(z)}{f_k'(0)} = \Big(\prod_{j=2}^{n} \frac{1}{|z_{k,j}|} \Big) \Big(\prod_{j=2}^{n} \frac{|z_{k,j}|}{z_{k,j}} \; \frac{z_{k,j} - z}{1 - \overline{z_{k,j}} z} \Big) z.$$

On the other hand, (6.5.3) reveals that

$$\begin{aligned} f_k'(0) &= \tau'_{f(a_k)}(f(a_k))\, f'(a_k)\, \tau'_{a_k}(0) \\ &= \frac{1 - |a_k|^2}{1 - |f(a_k)|^2} f'(a_k). \end{aligned}$$

This gives us the representation

$$f_k(z) = \frac{1 - |a_k|^2}{1 - |f(a_k)|^2} \;\; f'(a_k) \Big(\prod_{j=2}^{n} \frac{1}{|z_{k,j}|} \Big) \Big(\prod_{j=2}^{n} \frac{|z_{k,j}|}{z_{k,j}} \; \frac{z_{k,j} - z}{1 - \overline{z_{k,j}} z} \Big) z.$$

By (6.5.4) and a variation of Proposition 6.2.3,

$$\lim_{k \to \infty} \prod_{j=2}^{n} \frac{|z_{k,j}|}{z_{k,j}} \, \frac{z_{k,j} - z}{1 - \overline{z_{k,j}} \, z} \to 1$$

uniformly on compact subsets of $\mathbb{D}$. By Theorem 5.2.5,

$$\begin{aligned} \lim_{k \to \infty} \frac{1 - |f(a_k)|^2}{1 - |a_k|^2} &= \lim_{z \to \gamma} \frac{1 - |f(z)|^2}{1 - |z|^2} \\ &= |f'(\gamma)|. \end{aligned}$$

Therefore,

$$f_k(z) \to \frac{f'(\gamma)}{|f'(\gamma)|} z$$

uniformly on compact subsets of $\mathbb{D}$.

$(b) \implies (c)$ If (c) does not hold, then there is a sequence a_n in $\mathbb{D}$ that converges to a point $\zeta \in \mathbb{T}$ but

$$\lim_{k\to\infty} \frac{(1-|a_k|^2)|f'(a_k)|}{1-|f(a_k)|^2} \neq 1.$$

However, in the light of the formula

$$\frac{(1-|a_k|^2)f'(a_k)}{1-|f(a_k)|^2} = \big(\tau_{f(a_k)} \circ f \circ \tau_{a_k}\big)'(0),$$

we have a contradiction since $\tau_{f(a_k)} \circ f \circ \tau_{a_k}$ tends to a rotation and thus

$$\lim_{k\to\infty} |\big(\tau_{f(a_k)} \circ f \circ \tau_{a_k}\big)'(0)| = 1.$$

$(c) \implies (d)$ Since

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

one can see that the set of critical points $\{z \in \mathbb{D} : f'(z) = 0\}$ of f is finite. Toward a contradiction, assume that f has an asymptotic value $a \in \mathbb{D}$. By the Lindelöf theorem discussion earlier, f has a nontangential limit equal to a at $\zeta \in \mathbb{T}$. Consider the family

$$g_r := f \circ \phi_r, \quad 0 < r < 1,$$

where

$$\phi_r(z) = \frac{z + r\zeta}{1 + r\bar{\zeta} z}.$$

As $r \to 1$, the automorphisms ϕ_r converge uniformly on compact subsets of $\mathbb{D}$ to the constant function ζ. Therefore, g_r converges uniformly on compact subsets of $\mathbb{D}$ to the constant function a. Since $a \in \mathbb{D}$, we conclude that

$$\lim_{r\to 1} \frac{|g_r'(0)|}{1-|g_r(0)|^2} = 0.$$

The contradiction becomes more apparent when we note that

$$\frac{(1-|r\zeta|^2)\,|f'(r\zeta)|}{1-|f(r\zeta)|^2}=\frac{|g_r'(0)|}{1-|g_r(0)|^2},$$

and, by assumption, we must have

$$\lim_{r\to 1}\frac{(1-|r\zeta|^2)\,|f'(r\zeta)|}{1-|f(r\zeta)|^2}=1.$$

$(d) \implies (a)$ This is the bulk of the theorem and its most difficult part. We first show that condition (d) ($f \in \mathscr{S}$ has no asymptotic values in $\mathbb{D}$ and has a finite set of critical points) implies that f has finite constant valence. By Theorem 3.7.2 this will imply that f is a finite Blaschke product.

Suppose toward a contradiction that the valence function ν_f is not constant. Let

$$m=\min\{\nu_f(w) : w\in\mathbb{D}\}$$

and note that since ν_f is not constant, we have $0 \leqslant m < \infty$. Using the lower semicontinuity of v_f [73, Thm. 7.1] one can show that

$$\{w\in\mathbb{D} : \nu_f(w)=m\}$$

is closed in $\mathbb{D}$ and hence is a proper subset of $\mathbb{D}$.

One can also choose $b \in \mathbb{D}$ that belongs to the boundary of $\{\nu_f(w) = m : w \in \mathbb{D}\}$. Since f has a finite number of critical points, there is an $a \in \mathbb{D}$ such that the interval $[a, b)$ is free from any critical *values*, that is,

$$\{z\in\mathbb{D} : f'(z)=0\}\cap f^{-1}([a,b))=\varnothing \tag{6.5.5}$$

and $\nu_f(a) > m$ (b itself might be a critical value).

Now consider

$$\ell(t)=(1-t)a+tb,\quad 0\leqslant t\leqslant 1,$$

the line segment from a to b, and the analytic continuations of f^{-1} along ℓ. Starting with the initial point $c \in f^{-1}(a)$ (corresponding to $t = 0$), consider how far one of the analytic continuations of f^{-1} starting at c can go without running into difficulty in getting all the way to $t = 1$. We denote the curve (a "pullback") formed by this analytic continuation by γ_c. We need to verify that

(a) we can indeed get all the way to $t = 1$;
(b) γ_c does not approach $\mathbb{T}$.

Indeed, if c and d are distinct points in $f^{-1}(a)$, then γ_c and γ_d are necessarily disjoint paths in $\mathbb{D}$. Such paths can collide just at critical points which were already excluded by (6.5.5). Suppose that a pullback tends to $\mathbb{T}$. Then either

(a) it converges to a single point;
(b) it converges to a finite set of points on $\mathbb{T}$;
(c) it has an oscillatory nature and accumulates on a subarc $\mathbb{T}$.

Option (a) is excluded since f has no asymptotic values inside $\mathbb{D}$. The mere existence of such a curve that terminates at $\zeta \in \mathbb{T}$ means that $(1 - t_c)a + t_c b$ is an asymptotic value for f. Option (b) is also excluded since if the curve converges to a finite set of points on $\mathbb{T}$, then the curve has accumulation points in $\mathbb{D}$. This means that f is constant on these accumulation points, forcing f to be a constant function, which it is not. Option (c) is also excluded since otherwise, f would have the constant nontangential value $(1 - t_c)a + t_c b$ almost everywhere on this sub-arc which forces f to be a constant function. Here we are using a fact from the theory of H^∞ functions [38, Thm. 2.2] which says that if the non-tangential boundary values of an H^∞ function are equal to c almost everywhere on a subarc of $\mathbb{T}$, then this function is identically equal to the constant function c. Therefore all curves remain in $\mathbb{D}$ and since there are no critical values on $[a, b)$ we can pull back up to $t = 1$.

Now we arrive at our contradiction. Since $\nu_f(a) > m$, there are at least $m + 1$ such paths created above and all of them terminate at points of $f^{-1}(b)$. This forces $\nu_f(b) > m$, which contradicts the fact that $\{v_f(w) = m : w \in \mathbb{D}\}$ is closed and contains b on its boundary. Thus, ν_f is constant on $\mathbb{D}$.

We now show that ν_f is not identically equal to ∞ on $\mathbb{D}$. Suppose to the contrary that $\nu_f \equiv \infty$. First observe that the set of critical points is nonempty. If this was not the case, then f would be locally injective at each point of $\mathbb{D}$. Moreover, since f has no asymptotic values in $\mathbb{D}$, this means that the analytic continuation of f^{-1} between any two points of $\mathbb{D}$ remains within $\mathbb{D}$. By the monodromy theorem [100, Vol. III, Ch. 8], f^{-1} has an analytic continuation to all of $\mathbb{D}$. In other words, f is injective and thus $v_f \equiv 1$, which we are assuming is not the case.

Since f has a nonempty set of critical points, we let $w_1, w_2, \ldots, w_m$ denote the distinct critical values of f, in other words, $w_j = f(\xi_j)$ for some $\xi_j \in \mathbb{D}$ with $f'(\xi_j) = 0$. Consider m mutually disjoint smooth curves γ_k, parameterized by $t \in [0, 1]$, with

$$\gamma_k(0) = w_k, \quad \gamma_k([0, 1)) \subseteq \mathbb{D}, \quad \gamma_k(1) \in \mathbb{T}.$$

Let

$$\Omega = \mathbb{D} \setminus \bigcup_{k=1}^{m} \gamma_k([0, 1)).$$

Since Ω is an open, simply connected region, so is $f^{-1}(\Omega)$. We study the components of $f^{-1}(\Omega)$ to obtain a contradiction to our assumption that $v_f \equiv \infty$.

Let ω be a component of $f^{-1}(\Omega)$. Then the restricted function $f : \omega \to \Omega$ cannot have any asymptotic values in Ω. Otherwise, there is a curve in ω that tends to $\partial\omega$ and on this curve f tends to a limit inside Ω. By assumption, this curve cannot terminate on $\mathbb{T}$. Hence, it would have to terminate at some point $\zeta \in \mathbb{D} \cap \partial\omega$. Then $f(\zeta) \in \Omega$ and since ω is a component of $f^{-1}(\Omega)$ we must have $\zeta \in \omega$ which cannot be the case. In a manner similar to our previous discussion, the monodromy theorem implies that $f : \omega \to \Omega$ is a bijection.

Since $\nu_f \equiv \infty$, there are infinitely many components of $f^{-1}(\Omega)$. The set of critical points is finite and thus can meet the boundaries of only finitely many components. Hence, we may pick a component ω whose boundary does not contain of any critical point. Let $h : \Omega \to \omega$ be the inverse of $f : \omega \to \Omega$. Hence

$$f(h(z)) = z, \quad z \in \Omega. \tag{6.5.6}$$

We show that h can be analytically continued to all of $\mathbb{D}$.

Let us study the behavior of h as we approach $\gamma_k(t)$, for $0 \leqslant t < 1$. If Γ is any curve inside Ω that terminates at $\gamma_k(0)$, then the pullback $h^{-1}(\Gamma) \subseteq \omega$ cannot terminate at a point of $\mathbb{T}$. It also cannot oscillate toward a subarc of $\mathbb{T}$ (see the discussion above). Hence, by (6.5.6), it has to terminate at a point of $\mathbb{D} \cap \partial\omega$. By similar reasoning, oscillatory behavior is excluded even at the boundary of $\mathbb{D} \cap \partial\omega$. Moreover, for different curves Γ_1 and Γ_2 in Ω that terminate at $\gamma_k(0)$, their pullbacks cannot converge to different points of $\mathbb{D} \cap \partial\omega$. By (6.5.6), the injectivity of $f : \omega \to \Omega$ would be violated. This means that

$$\zeta_k = \lim_{\substack{z \to \gamma_k(0) \\ z \in \Omega}} h(z)$$

exists and belongs to $\mathbb{D} \cap \partial\omega$. In particular, $f'(\zeta_k) \neq 0$.

Using a similar argument, h has a limit when we approach $\gamma_k(t)$, for $0 < t < 1$, from one side of the arc γ_k. Hence one-sided limits exist at all points $\gamma_k(t)$, for $0 < t < 1$, and h is necessarily continuous on each side and its limiting values converge to ζ_k when we move toward $\gamma_k(0)$ from either side. Since f is univalent on a small neighborhood of ζ_k and $f(\zeta_k) = \gamma_k(0)$, we see that h coincides with f^{-1} on a small neighborhood of $\gamma_k(0)$. Therefore, the one-sided limits of h must agree at least for small values of $t > 0$. Thus, h has continuous extension to

$$\Omega \cup \{\gamma_k(t) : 0 \leqslant t < \tau\}$$

for some value of $\tau \in (0, 1)$. By Morera's theorem, it has an analytic continuation to $\Omega \cup \{\gamma_k(t) : 0 \leqslant t < \tau\}$. As a matter of fact, we must have $\sup \tau = 1$. Otherwise, we could repeat the argument above with $\gamma_k(\tau)$ playing the role of $\gamma_k(0)$ and extend further and thus obtain a contradiction. In short, h extends analytically to $\Omega \cup \{\gamma_k(t) : 0 \leqslant t < 1\}$, and hence to $\mathbb{D}$. Let us denote this extension by H. According to (6.5.6), we have

$$f(H(z)) = z, \quad z \in \mathbb{D}.$$

Set $a = H(0)$ and write the identity above as

$$(f \circ \tau_a) \circ (\tau_a \circ H)(z) = z, \quad z \in \mathbb{D}. \tag{6.5.7}$$

Since f is not a conformal mapping (this was ruled out at the beginning), the Schwarz Lemma says that

$$|(f \circ \tau_a)'(0)| < 1. \tag{6.5.8}$$

Also by the Schwarz Lemma,

$$|(\tau_a \circ H)'(0)| \leqslant 1, \tag{6.5.9}$$

since $(\tau_a \circ H)(0) = 0$ and $\tau_a \circ H$ maps $\mathbb{D}$ into itself. However, (6.5.7) implies that

$$(f \circ \tau_a)'(0) \times (\tau_a \circ H)'(0) = 1,$$

which is a contradiction of (6.5.8) and (6.5.9). The proof is now complete. □

6.6 Notes

Critical Values

For a finite Blaschke product B of degree n,

$$\{w \in \mathbb{D} : w = B(z),\ B'(z) = 0\}$$

is the set of *critical values* of B. We state the following result from [6]. For each critical value w_j, there are most n distinct points in $B^{-1}(\{w_j\})$. The number

$$\delta_B(w_j) = n - |B^{-1}(\{w_j\})|,$$

where $|E|$ is the cardinality of a set E, is the *deficiency* of B at w_j. One can show that if $w_1, w_2, \ldots, w_k$ are the critical values of B, then

$$\sum_{j=1}^{k} \delta_B(w_j) = n - 1.$$

Moreover, for distinct points $w_1, w_2, \ldots, w_k \in \mathbb{D}$ and $\delta_1, \delta_2, \ldots, \delta_k \in \mathbb{N}$ such that $\sum_{j=1}^{k} \delta_j = n-1$, there is a finite Blaschke product of degree n whose critical values are $\{w_1, w_2, \ldots, w_k\}$ and with $\delta_B(w_j) = \delta_j$ for all $j = 1, 2, \ldots, k$. The paper [6] also discusses a version of this theorem for polynomials.

Asymptotic Values

There is a literature concerning asymptotic values of functions (continuous, harmonic, analytic) on $\mathbb{D}$ (recall Definition 6.5.1) [4, 25, 84, 85, 102].

6.7 Exercises

6.1 Show that if P is a polynomial of degree two, then the zero of P' is the average of the roots of P.

6.2 Show that if P is a polynomial of degree three with distinct zeros, then the zeros of P' are the foci of the ellipse that is tangent to the midpoints of the triangle determined by the zeros of P. This is known as *Marden's theorem* and the ellipse is known as the *Steiner inellipse*. See [86] for a history of this theorem as well as a proof.

6.3 Let

$$B(z) = \left(\frac{a-z}{1-\overline{a}z}\right)^m \left(\frac{b-z}{1-\overline{b}z}\right)^n \left(\frac{c-z}{1-\overline{c}z}\right)^p,$$

in which $a, b, c \in \mathbb{D}$ and $m, n, p \geqslant 1$. Find the zeros of B' and show that as m, n, p range over the positive integers, the zeros of B' form a dense subset of the hyperbolic convex hull of a, b, c.

6.4 Show that an entire function is proper if and only if it is a polynomial.
Hint: Use Picard's theorem.

6.5 Show that an analytic function $f : \mathbb{D} \to \mathbb{D}$ is proper if and only if it is a finite Blaschke product.
Hint: Use Theorem 3.5.2.

Chapter 7
Interpolation

Interpolation of data by functions from a given class has a rich history dating back to Newton and Lagrange. Famous examples involve interpolation by polynomials, rational functions, and bounded analytic functions. This chapter covers various types of interpolation and the connection these problems make to finite Blaschke products.

In a typical interpolation problem, one considers a class of analytic functions $\mathscr{F}$ on a domain $\Omega \subseteq \mathbb{C}$ and asks, for a given list of distinct points $z_1, z_2, \ldots, z_n \in \Omega$ and a given list of values $w_1, w_2, \ldots, w_n \in \mathbb{C}$, if there is an $f \in \mathscr{F}$ such that

$$f(z_k) = w_k, \qquad 1 \leqslant k \leqslant n. \tag{7.0.1}$$

There are additional questions that can be considered.

(i) Characterize all pairs $(z_1, w_1), (z_2, w_2), \ldots, (z_n, w_n)$ for which (7.0.1) has a solution in $\mathscr{F}$.
(ii) Characterize the set of points $z_1, z_2, \ldots, z_n$ such that (7.0.1) has a solution in $\mathscr{F}$ for all values of $w_1, w_2, \ldots, w_n$ in a fixed given set.
(iii) If the interpolation problem has a solution, give an explicit formula for it or provide an algorithm to find it.
(iv) If the interpolation problem has a solution, is the solution unique? If the solution is not unique, find one that is extremal with respect to a given property.

In this chapter, we study interpolation by finite Blaschke products. We focus on the following two questions.

(a) For distinct $z_1, z_2, \ldots, z_n \in \mathbb{D}$ and any $w_1, w_2, \ldots, w_n \in \mathbb{D}$, is there a finite Blaschke product B such that $B(z_k) = w_k$ for $1 \leqslant k \leqslant n$?
(b) For distinct $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ and any $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{T}$, is there a finite Blaschke product B such that $B(\zeta_k) = \xi_k$ for $1 \leqslant k \leqslant n$?

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_7

The first of these problems (interpolation in $\mathbb{D}$) was settled long ago by Pick. The second problem (boundary interpolation), although settled, is more delicate. We give several solutions to these problems. Some of these are short existence proofs while others are longer but more constructive.

7.1 Lagrange Interpolation: Polynomials

In order to place this subject in context, we begin with some classical interpolation results. When $\mathscr{F}$ is the set of analytic polynomials

$$\Big\{\sum_{k=0}^{n} a_k z^k : a_k \in \mathbb{C}, n \geqslant 0\Big\},$$

the following classical result settles the interpolation problem (7.0.1).

Theorem 7.1.1 (Lagrange Interpolation Theorem) *Given distinct* $z_1, z_2, \ldots, z_n \in \mathbb{C}$ *and any* $w_1, w_2, \ldots, w_n \in \mathbb{C}$*, there is a unique polynomial* P *of degree at most* $n-1$ *such that*

$$P(z_k) = w_k, \qquad 1 \leqslant k \leqslant n. \tag{7.1.2}$$

Proof For $k = 1, 2, \ldots, n$, define the *Lagrange polynomials* by

$$L_k(z) = \prod_{\substack{i=1\\ i\neq k}}^{n} \frac{z - z_i}{z_k - z_i}, \qquad 1 \leqslant k \leqslant n, \tag{7.1.3}$$

and verify that $\deg L_k = n-1$ and $L_k(z_j) = \delta_{jk}$, the Kronecker delta function. The polynomial

$$P = \sum_{k=1}^{n} w_k L_k$$

is of degree at most $n-1$ and satisfies $P(z_k) = w_k$ for $1 \leqslant k \leqslant n$. If Q is another solution to the interpolation problem (7.1.2) and $\deg Q \leqslant n-1$, then $P-Q$ vanishes at the n distinct points $z_1, z_2, \ldots, z_n$ and is of degree at most $n-1$. Thus, $P-Q \equiv 0$ and hence $P = Q$. Therefore, P is the unique solution to (7.1.2) of degree at most $n-1$. □

Observe that

$$Q(z) = (z-z_1)\cdots(z-z_n) \implies L_k(z) = \frac{Q(z)}{Q'(z_k)(z-z_k)}, \qquad 1 \leqslant k \leqslant n.$$

See Exercise 7.1 for another proof of Theorem 7.1.1 and see Exercises 7.2, 7.3, and 7.4 for further applications. There is also *Hermite interpolation* [87] which interpolates not only the function but also its derivatives; see the notes at the end of this chapter.

7.2 Lagrange Interpolation: Rational Functions

In this section, we treat some interpolation problems involving rational functions. These results will be used later on for interpolation by finite Blaschke products. In what follows,

$$\widehat{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$$

denotes the extended real line (as a subset of the Riemann sphere $\widehat{\mathbb{C}}$).

Lemma 7.2.1 (Gorkin–Rhoades [65]) *Let f be the rational function*

$$f(z) = \frac{(z - x_1)(z - x_2) \cdots (z - x_n)}{(z - p_1)(z - p_2) \cdots (z - p_n)},$$

where $x_1, x_2, \ldots, x_n$ and $p_1, p_2, \ldots, p_n$ are real numbers for which

$$p_1 < x_1 < p_2 < x_2 < \cdots < p_n < x_n. \tag{7.2.2}$$

Then f satisfies the following.

(a) $f(\mathbb{C}_+) \subseteq \mathbb{C}_+$.
(b) $f(\mathbb{C}_-) \subseteq \mathbb{C}_-$.
(c) $f(\widehat{\mathbb{R}}) \subseteq \widehat{\mathbb{R}}$.
(d) f has a simple zero at each $x_1, x_2, \ldots, x_n$.
(e) f has a simple pole with a negative residue at each $p_1, p_2, \ldots, p_n$.

Proof We leave it to the reader to verify (c), (d) and the first part of (e). To prove the second part of (e), perform a partial fraction expansion, noting that f is the quotient of two monic polynomials of equal degree, along with the fact that

$$\lim_{z \to \infty} f(z) = 1,$$

to obtain

$$f(z) = 1 + \frac{\lambda_1}{z - p_1} + \cdots + \frac{\lambda_n}{z - p_n}. \tag{7.2.3}$$

Fix $1 \leqslant k \leqslant n$ and multiply both sides of the previous equation by $z - p_k$ and then set $z = p_k$ to see that

$$\lambda_k = \frac{\displaystyle\prod_{j=1}^{n}(p_k - x_j)}{\displaystyle\prod_{\substack{j=1 \\ j \neq k}}^{n}(p_k - p_j)}.$$

Rewrite this as

$$\lambda_k = (p_k - x_k) \cdot \frac{\displaystyle\prod_{j=1}^{k-1}(p_k - x_j)}{\displaystyle\prod_{j=1}^{k-1}(p_k - p_j)} \cdot \frac{\displaystyle\prod_{j=k+1}^{n}(p_k - x_j)}{\displaystyle\prod_{j=k+1}^{n}(p_k - p_j)}$$

and observe, via the hypothesis (7.2.2), that $\lambda_k < 0$.

If $\alpha < 0$ and $\beta \in \mathbb{R}$, the function

$$g(z) = \frac{\alpha}{z - \beta}$$

satisfies

$$\operatorname{Im} g(z) = -\alpha \frac{\operatorname{Im}(z)}{|z - \beta|^2}$$

and hence g satisfies (a) and (b). By (7.2.3), the function $f - 1$ is a finite sum of such functions and hence must also satisfy (a) and (b). □

The function

$$f(z) = a + \frac{b}{z - c}, \tag{7.2.4}$$

in which $a, b, c \in \mathbb{R}$ and $b < 0$, is from a class of functions introduced in Lemma 7.2.1. We will encounter this function in the proof of Corollary 7.2.6.

The following result resembles Lagrange interpolation. A weaker version of this result was first given by Younis [138]. Even though the method of Younis was constructive, the degree of his interpolating function could be as large as $n^2 - n$; see Exercise 7.5. The following provides a constructive solution of degree n. Moreover, the solution exhibits further interesting properties.

Theorem 7.2.5 (Gorkin–Rhoades [65]) *Let $x_1, x_2, \ldots, x_n$ and $p_1, p_2, \ldots, p_n$ be real numbers that satisfy*

$$p_1 < x_1 < p_2 < x_2 < \cdots < p_n < x_n$$

and let $y_1, y_2, \ldots, y_n$ be any real numbers. Then there is a rational function f of degree n that satisfies the following.

(a) $f(\mathbb{C}_+) \subseteq \mathbb{C}_+$.
(b) $f(\mathbb{C}_-) \subseteq \mathbb{C}_-$.
(c) $f(\widehat{\mathbb{R}}) \subseteq \widehat{\mathbb{R}}$.
(d) $f(x_k) = y_k$ *for* $1 \leqslant k \leqslant n$.
(e) f *has a simple pole with a negative residue at each* $p_1, p_2, \ldots, p_n$.

Proof Without loss of generality, we can assume that $y_1, y_2, \ldots, y_n > 0$. If this is not the case, let $M \in \mathbb{R}$ be so large that $y_k' = M + y_k > 0$ for $1 \leqslant k \leqslant n$. Then solve the interpolation problem for $g(x_k) = y_k'$ as described below. The answer to the original interpolation problem is $f = g - M$.

The functions

$$f_k(z) = \prod_{\substack{j=1\\ j\neq k}}^{n} \frac{z - x_j}{z - p_j}, \qquad 1 \leqslant k \leqslant n,$$

satisfy the properties described in Lemma 7.2.1. Moreover, for each k,

$$\begin{aligned} f_k(x_k) &= \prod_{\substack{j=1\\ j\neq k}}^{n} \frac{x_k - x_j}{x_k - p_j} \\ &= \prod_{j=1}^{k-1} \frac{x_k - x_j}{x_k - p_j} \cdot \prod_{j=k+1}^{n} \frac{x_k - x_j}{x_k - p_j} > 0 \end{aligned}$$

and $f_k(x_j) = 0$ for $j \neq k$. Define

$$f = \sum_{k=1}^{n} \frac{y_k}{f_k(x_k)}\, f_k.$$

Since the coefficients $y_k / f_k(x_k)$ are all positive and since the f_k satisfy the properties described in Lemma 7.2.1, f satisfies properties (a), (b), and (c). The coefficients were chosen so that $f(x_k) = y_k$.

The function f is of degree at most n and its possible poles are $p_1, p_2, \ldots, p_n$. To ensure that these singularities are not removable, condition (e) of Lemma 7.2.1 implies that

$$\operatorname{Res}(f, p_k) = \sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{y_j}{f_j(x_j)} \operatorname{Res}(f_j, p_k) < 0.$$

Hence the degree of f is precisely n. This completes the proof. □

The following important corollary can be used to obtain theorems about boundary interpolation by finite Blaschke products; see Sect. 7.5.

Corollary 7.2.6 *Let $x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $a_1, a_2, \ldots, a_n \in \mathbb{R}$ satisfy*

$$a_1 < x_1 < a_2 < x_2 < \cdots < a_n < x_n$$

and let $y_0, y_1, \ldots, y_n \in \mathbb{R}$ satisfy $y_0 \neq y_k$ for $1 \leqslant k \leqslant n$. Then there is a rational function f of degree n that satisfies the following.

(a) $f(\mathbb{C}_+) \subseteq \mathbb{C}_+$.
(b) $f(\mathbb{C}_-) \subseteq \mathbb{C}_-$.
(c) $f(\widehat{\mathbb{R}}) \subseteq \widehat{\mathbb{R}}$.
(d) $f(x_k) = y_k$ *for* $1 \leqslant k \leqslant n$.
(e) $f(a_k) = y_0$ *for* $1 \leqslant k \leqslant n$.

Proof Let

$$g(z) = y_0 - \frac{1}{z}.$$

By (7.2.4), g satisfies the properties in Lemma 7.2.1. Let

$$y_k' = g^{-1}(y_k), \qquad 1 \leqslant k \leqslant n.$$

Since $y_0 \neq y_k$, we have $y_k' \in \mathbb{R}$ (only y_0 is sent to $y_0' = \infty$). Theorem 7.2.5 says that there is a rational function h of degree n that satisfies (a), (b), and (c), and such that

$$h(x_k) = y_k'$$

with simple poles at points a_k for $1 \leqslant k \leqslant n$.

Set $f = g \circ h$. Then f is a rational function of degree n that satisfies (a), (b), and (c), and such that

$$f(x_k) = g(h(x_k)) = g(y_k') = y_k, \qquad 1 \leqslant k \leqslant n.$$

Moreover, since h has a simple pole at each a_k,

$$f(a_k) = g(h(a_k)) = g(\infty) = y_0, \qquad 1 \leqslant k \leqslant n.$$

Thus, f is the desired function. □

7.3 Pick Interpolation Theorem

We now discuss a famous interpolation problem for the Schur class $\mathscr{S}$. Given distinct $z_1, z_2, \dots, z_n \in \mathbb{D}$ and arbitrary $w_1, w_2, \dots, w_n \in \mathbb{C}$, is there an $f \in \mathscr{S}$ such that

$$f(z_k) = w_k, \qquad 1 \leqslant k \leqslant n?$$

Since $f \in \mathscr{S}$, if there is a solution to this problem we must have $|w_j| \leqslant 1$ for all j. Moreover, if $w_{j_0} \in \mathbb{T}$ for some $j_0 \in \{1, 2, \dots, n\}$, then by the Maximum Modulus Principle, $f \equiv w_{j_0}$. Even when all the z_j and w_j belong to $\mathbb{D}$, we cannot always solve interpolation problem. For example, if $f \in \mathscr{S}$ solves the two-point interpolation problem

$$f(z_1) = w_1 \quad \text{and} \quad f(0) = 0, \tag{7.3.1}$$

then the Schwarz Lemma (Lemma 1.1.1) tells us that $|w_1| = |f(z_1)| \leqslant |z_1|$. Thus, a necessary condition for the solvability of (7.3.1) is

$$|w_1| \leqslant |z_1|. \tag{7.3.2}$$

A closer examination of the generic two-point interpolation problem

$$f(z_1) = w_1 \quad \text{and} \quad f(z_2) = w_2 \tag{7.3.3}$$

is instructive, although it requires a few important facts from linear algebra. Recall that for vectors

$$\mathbf{x} = (x_1, \dots, x_n), \quad \mathbf{y} = (y_1, \dots, y_n) \in \mathbb{C}^n,$$

their inner product $\langle \mathbf{x}, \mathbf{y} \rangle$ is

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j \overline{y_j}.$$

If M_n denotes the set of all $n \times n$ complex matrices and $A \in M_n$, then A^* denotes the conjugate transpose of A and $\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^*\mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$.

Definition 7.3.4 $A \in M_n$ is *positive semidefinite* if

$$\langle A\mathbf{x}, \mathbf{x} \rangle \geqslant 0, \quad \mathbf{x} \in \mathbb{C}^n.$$

A positive semidefinite matrix A is automatically Hermitian: $A = A^*$. The spectral theorem says that a Hermitian matrix is positive semidefinite if and only if its eigenvalues are nonnegative; see Exercise 7.6. If $A \in M_n$ is positive semidefinite and S is $m \times n$, then SAS^* is positive semidefinite; see Exercise 7.7.

Returning to our discussion of the two point interpolation problem (7.3.3), the Schwarz–Pick theorem (Theorem 1.4.1) tells us that

$$\left|\frac{f(z_1)-f(z_2)}{1-\overline{f(z_2)}f(z_1)}\right| \leqslant \left|\frac{z_1-z_2}{1-\overline{z_2}z_1}\right|, \qquad z_1, z_2 \in \mathbb{D}$$

for each $f \in \mathscr{S}$. If $f \in \mathscr{S}$ satisfies (7.3.3), then (1.6.1) yields

$$\begin{aligned}
\left|\frac{w_1-w_2}{1-\overline{w_2}w_1}\right| \leqslant \left|\frac{z_1-z_2}{1-\overline{z_2}z_1}\right| &\iff 1-\left|\frac{z_1-z_2}{1-\overline{z_2}z_1}\right|^2 \leqslant 1-\left|\frac{w_1-w_2}{1-\overline{w_2}w_1}\right|^2 \\
&\iff \frac{(1-|z_1|^2)(1-|z_2|^2)}{|1-\overline{z_2}z_1|^2} \leqslant \frac{(1-|w_1|^2)(1-|w_2|^2)}{|1-\overline{w_2}w_1|^2} \\
&\iff \frac{|1-\overline{w_2}w_1|^2}{|1-\overline{z_2}z_1|^2} \leqslant \frac{(1-|w_1|^2)(1-|w_2|^2)}{(1-|z_1|^2)(1-|z_2|^2)} \\
&\iff \det\begin{bmatrix} \dfrac{1-|w_1|^2}{1-|z_1|^2} & \dfrac{1-w_1\overline{w_2}}{1-z_1\overline{z_2}} \\ \dfrac{1-\overline{w_1}w_2}{1-\overline{z_1}z_2} & \dfrac{1-|w_2|^2}{1-|z_2|^2} \end{bmatrix} \geqslant 0.
\end{aligned}$$

Since $z_1, z_2, w_1, w_2 \in \mathbb{D}$, the trace of the Hermitian matrix

$$P(z_1, z_2; w_1, w_2) = \begin{bmatrix} \dfrac{1-|w_1|^2}{1-|z_1|^2} & \dfrac{1-w_1\overline{w_2}}{1-z_1\overline{z_2}} \\ \dfrac{1-\overline{w_1}w_2}{1-\overline{z_1}z_2} & \dfrac{1-|w_2|^2}{1-|z_2|^2} \end{bmatrix} \tag{7.3.5}$$

is positive and thus the sum of its eigenvalues is positive. Since the determinant of a square matrix equals the product of its eigenvalues,

$$\det P(z_1, z_2; w_1, w_2) \geqslant 0$$

if and only if (7.3.5) is positive semidefinite. Consequently, a necessary condition for the solvability of the two-point interpolation problem (7.3.3) is the positive semidefiniteness of (7.3.5). If $z_2 = w_2 = 0$, then

$$0 \leqslant \det P(z_1, 0; w_1, 0) = \frac{1-|w_1|^2}{1-|z_1|^2} - 1 \qquad \iff \qquad |w_1| \leqslant |z_1|,$$

which recovers the necessary condition (7.3.2) for the solvability of the special two-point problem (7.3.1).

The following celebrated result of Pick provides a complete solution to the n-point interpolation problem. It has been extended in many directions (see [1] for a thorough discussion).

Theorem 7.3.6 (Pick [114]) *Suppose*

$$z_1, z_2, \ldots, z_n \in \mathbb{D}$$

are distinct and $w_1, w_2 \ldots, w_n \in \mathbb{C}$. *There is an* $f \in \mathscr{S}$ *such that*

$$f(z_k) = w_k, \qquad 1 \leqslant k \leqslant n, \tag{7.3.7}$$

if and only if

$$P = P(z_1, \ldots, z_n; w_1, \ldots, w_n) = \left[\frac{1 - w_i \overline{w_j}}{1 - z_i \overline{z_j}}\right]_{i,j=1}^{n} \tag{7.3.8}$$

is positive semidefinite. Moreover, if a solution exists, then there is a finite Blaschke product of degree at most n that does the interpolation in (7.3.7).

Remark 7.1 Before proceeding to the proof, we make a few remarks about some degenerate situations that might occur when applying an induction argument below.

(a) If $|w_{j_0}| > 1$ for some $j_0 \in \{1, 2, \ldots, n\}$, then, since $|f| \leqslant 1$ for all $f \in \mathscr{S}$, there is no solution to the interpolation problem (7.3.7). Moreover, still under the assumption that $|w_{j_0}| > 1$ for some j_0, the corresponding Pick matrix has the property that $\langle P\mathbf{e}_{j_0}, \mathbf{e}_{j_0}\rangle = 1 - |w_{j_0}|^2 < 0$.
(b) If $|w_{j_0}| = 1$ for some $j_0 \in \{1, 2, \ldots, n\}$, then for the interpolation problem (7.3.7) to have a solution, it must be the case that $f \equiv w_{j_0}$ and consequently $w_j = w_{j_0}$ for all j. Furthermore, this constant function solution to the interpolation is a finite Blaschke product of degree 0. The Pick matrix P is the zero matrix, which is positive semidefinite.
(c) If $z_0 = w_0 = 0$ and f is a solution to the interpolation problem (7.3.7), then the Schwarz lemma says that either $|f(z_j)| < |w_j|$ for all $j \in \{2, 3, \ldots, n\}$ or there exists a $\xi \in \mathbb{T}$ such that $f(z) = \xi z$ (and consequently $z_j = \xi w_j$ for all j).

Proof (of Theorem 7.3.6) As mentioned in Remark 7.1, we can assume that $|w_j| \leqslant 1$ for all j. We proceed by induction on n. The base case is $n = 1$. If $|w_1| = 1$, then the constant function $f \equiv w_1$, which is a finite Blaschke product of degree 0, accomplishes the interpolation. Furthermore, the associated 1×1 (scalar) Pick matrix is

$$P(z_1, w_1) = \frac{1 - |w_1|^2}{1 - |z_1|^2} = 0,$$

which is positive semidefinite. If $|w_1| < 1$, the one-point interpolation problem $f(z_1) = w_1$ has the solution $f = \tau_{w_1} \circ \tau_{z_1}$, which is a finite Blaschke product of degree 1. Since the 1×1 (scalar) Pick matrix

$$P(z_1, w_1) = \frac{1 - |w_1|^2}{1 - |z_1|^2}$$

is positive semidefinite for all $z_1, w_1 \in \mathbb{D}$, the theorem is true when $n = 1$.

For our induction hypothesis, suppose that the theorem holds for $n - 1$ points. First observe that the n-point interpolation problem (7.3.7) has a solution with $f \in \mathscr{S}$ that is a finite Blaschke product of degree at most n if and only if, for any $w_0, z_0 \in \mathbb{D}$, the interpolation problem

$$g(\tau_{z_0}(z_k)) = \tau_{w_0}(w_k), \qquad 1 \leqslant k \leqslant n, \tag{7.3.9}$$

has a solution with $g \in \mathscr{S}$ that is a finite Blaschke product of degree at most n. Indeed, write

$$f = \tau_{w_0} \circ g \circ \tau_{z_0} \tag{7.3.10}$$

in order to pass from a solution to (7.3.7) to a solution to (7.3.9) and back; Lemma 3.6.1 says that $\deg f = \deg g$. Now observe that the identity

$$\frac{1 - \tau_{w_0}(w_i)\overline{\tau_{w_0}(w_j)}}{1 - \tau_{z_0}(z_i)\overline{\tau_{z_0}(z_j)}} = \frac{1 - |w_0|^2}{1 - |z_0|^2}\,\frac{1 - \overline{w_0}w_i}{1 - \overline{z_0}z_i} \cdot \frac{1 - w_i\overline{w_j}}{1 - z_i\overline{z_j}} \cdot \frac{1 - w_0\overline{w_j}}{1 - z_0\overline{z_j}} \tag{7.3.11}$$

implies that

$$\left[\frac{1 - \tau_{w_0}(w_i)\overline{\tau_{w_0}(w_j)}}{1 - \tau_{z_0}(z_i)\,\overline{\tau_{z_0}(z_j)}}\right]_{i,j=1}^{n} = \frac{1 - |w_0|^2}{1 - |z_0|^2}\Lambda P \Lambda^*, \tag{7.3.12}$$

in which Λ is the $n \times n$ diagonal matrix

$$\Lambda = \operatorname{diag}\left(\frac{1 - \overline{w_0}w_1}{1 - \overline{z_0}z_1}, \frac{1 - \overline{w_0}w_2}{1 - \overline{z_0}z_2}, \ldots, \frac{1 - \overline{w_0}w_n}{1 - \overline{z_0}z_n}\right);$$

see Exercise 7.8. Therefore, P is positive semidefinite if and only if

$$\left[\frac{1 - \tau_{w_0}(w_i)\,\overline{\tau_{w_0}(w_j)}}{1 - \tau_{z_0}(z_i)\,\overline{\tau_{z_0}(z_j)}}\right]_{i,j=1}^{n}$$

is positive semidefinite; see Exercise 7.7. For simplicity in our labeling, set

$$z_j' = \tau_{z_n}(z_j), \quad w_j' = \tau_{w_n}(w_j)$$

and then relabel so that $z_j = z'_j$ and $w_j = w'_j$. With this relabeling we have $z_n = w_n = 0$ (this is important to remember in practice when we want to find an explicit solution to the interpolation problem). Then each entry of the last row and column of

$$A = \left[\frac{1 - w_i\overline{w_j}}{1 - z_i\overline{z_j}}\right]_{i,j=1}^{n}$$

is 1. Let S be the following $n \times n$ matrix

$$S = \begin{bmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

which is invertible, and observe that

$$SAS^* = \begin{bmatrix} \left[\frac{1-w_i\overline{w_j}}{1-z_i\overline{z_j}} - 1\right]_{i,j=1}^{n-1} & [0]_{(n-1)\times 1} \\ [0]_{1\times(n-1)} & 1 \end{bmatrix}. \tag{7.3.13}$$

Consequently, A is positive semidefinite if and only if

$$\left[\frac{1 - w_i\overline{w_j}}{1 - z_i\overline{z_j}} - 1\right]_{i,j=1}^{n-1}$$

is positive semidefinite. Since

$$\left[\frac{1 - w_i\overline{w_j}}{1 - z_i\overline{z_j}} - 1\right]_{i,j=1}^{n-1} = D\left[\frac{1 - \frac{w_i}{z_i}\overline{\frac{w_j}{z_j}}}{1 - z_i\overline{z_j}}\right]_{i,j=1}^{n-1} D^*, \tag{7.3.14}$$

where

$$D = \operatorname{diag}(z_1, z_2, \ldots, z_{n-1}),$$

the matrix

$$\left[\frac{1 - w_i\overline{w_j}}{1 - z_i\overline{z_j}}\right]_{i,j=1}^{n},$$

in which $z_n = w_n = 0$, is positive semidefinite if and only if the $(n-1)\times(n-1)$ matrix

$$\left[\frac{1 - \frac{w_i}{z_i}\overline{\frac{w_j}{z_j}}}{1 - z_i\overline{z_j}}\right]_{i,j=1}^{n-1}$$

is positive semidefinite (recall that $z_1, z_2, \ldots, z_{n-1}$ are nonzero since $z_n = 0$ and the z_j are distinct). By the induction hypothesis, this occurs if and only if there is a finite Blaschke product g of degree at most $n-1$ that solves the $(n-1)$-point interpolation problem

$$g(z_i) = \frac{w_i}{z_i}, \qquad 1 \leqslant i \leqslant n-1. \tag{7.3.15}$$

If $f \in \mathscr{S}$ satisfies the n-point interpolation problem

$$f(z_k) = w_k, \quad 1 \leqslant k \leqslant n-1, \quad \text{and} \quad f(0) = 0, \tag{7.3.16}$$

then $g(z) = f(z)/z$ belongs to $\mathscr{S}$ (Schwarz Lemma) and is a solution to the $(n-1)$-point interpolation problem (7.3.15). Conversely if $g \in \mathscr{S}$ solves (7.3.15), then

$$f(z) = zg(z) \tag{7.3.17}$$

solves (7.3.16). See Remark 7.1 about what happens if $|w_j/z_j| = 1$ for some j. Also observe that g is a finite Blaschke product of degree $n-1$ if and only if f is a finite Blaschke product of degree n. This completes the induction. □

Among its many consequences, the preceding theorem can be used to provide a new proof of the Carathéodory approximation theorem (Theorem 4.1.1); see Exercise 7.12.

We now discuss the uniqueness of the solution to the Pick problem.

Theorem 7.3.18 *Suppose*

$$z_1, z_2, \ldots, z_n \in \mathbb{D}$$

are distinct and $w_1, w_2 \ldots, w_n \in \mathbb{C}$. *If*

$$P = \left[\frac{1 - w_i\overline{w_j}}{1 - z_i\overline{z_j}}\right]_{i,j=1}^{n}$$

is positive semidefinite, then the n*-point interpolation problem*

$$f(z_k) = w_k, \qquad 1 \leqslant k \leqslant n, \tag{7.3.19}$$

in which $f \in \mathscr{S}$, *has a unique solution if and only if* $\det P = 0$. *In this case,* f *is a finite Blaschke product of degree* $m = \operatorname{rank} P$. *Conversely, if* (7.3.19) *is satisfied by a finite Blaschke product of degree* $m < n$, *then the solution is unique and* $m = \operatorname{rank} P$.

Proof As in the proof of Theorem 7.3.6, we may assume that $z_n = w_n = 0$. Then (7.3.17) shows that (7.3.19) is uniquely solved by a finite Blaschke product of degree

m if and only if the corresponding $(n-1)$-point problem (7.3.15) for g is uniquely solved by a finite Blaschke product of degree $m-1$. Moreover, (7.3.13) and (7.3.14) show that

$$\operatorname{rank} P_n = 1 + \operatorname{rank} P_{n-1},$$

in which we employ the notation of (7.3.8), namely

$$P_n = P(z_1, \ldots, z_n; w_1, \ldots, w_n)$$

and

$$P_{n-1} = P(z_1, \ldots, z_{n-1}; \tfrac{w_1}{z_1}, \ldots, \tfrac{w_{n-1}}{z_{n-1}}).$$

The result now follows by induction.

If the interpolation problem is satisfied by a finite Blaschke product of degree $m < n$, then after m repetitions of the procedure above, the remaining interpolation problem must be solved by a finite Blaschke product of degree zero; that is, a unique constant unimodular function. Thus, the corresponding matrix for the final interpolation problem is identically zero, which shows that the solution is unique and rank $P = m$. $\square$

Let

$$P = \left[\frac{1 - w_i \overline{w_j}}{1 - z_i \overline{z_j}}\right]_{i,j=1}^{n},$$

in which $z_1, z_2, \ldots, z_n \in \mathbb{D}$ are distinct and $w_1, w_2, \ldots, w_n \in \mathbb{C}$ are arbitrary. Theorem 7.3.18 tells us that if rank $P = m < n$ and if we choose $m+1$ indices $k_1, k_2, \ldots, k_{m+1}$ among $1, 2, \ldots, n$ such that

$$\det\left[\frac{1 - w_{k_i} \overline{w_{k_j}}}{1 - z_{k_i} \overline{z_{k_j}}}\right]_{i,j=1}^{m} > 0 \quad \text{and} \quad \det\left[\frac{1 - w_{k_i} \overline{w_{k_j}}}{1 - z_{k_i} \overline{z_{k_j}}}\right]_{i,j=1}^{m+1} = 0,$$

then the conditions

$$f(z_{k_i}) = w_{k_i}, \qquad 1 \leqslant i \leqslant m+1,$$

determine a unique finite Blaschke product f of degree m. Moreover, this finite Blaschke product automatically satisfies $f(z_k) = w_k$ for those indices k that are not among $k_1, k_2, \ldots, k_{m+1}$. Although this may seem surprising at first, one should keep in mind that the hypothesis that rank $P = m < n$ says that the remaining conditions are dependent on the data $z_{k_1}, z_{k_2}, \ldots, z_{k_{m+1}}$ and $w_{k_1}, w_{k_2}, \ldots, w_{k_{m+1}}$.

If $\det P > 0$, which is equivalent to rank $P = n$, then there are infinitely many solutions. Indeed, if $n = 1$ and $z_1 = w_1 = 0$, then P is the 1×1 scalar matrix $[1]$ and any $f \in \mathscr{S}$ with $f(0) = 0$ is a solution. In particular, any finite Blaschke product that vanishes at the origin is a solution.

Corollary 7.3.20 *Suppose*

$$z_1, z_2, \ldots, z_n \in \mathbb{D}$$

are distinct and

$$w_1, w_2 \ldots, w_n \in \mathbb{D}.$$

Suppose that

$$P = \left[\frac{1 - w_i \overline{w_j}}{1 - z_i \overline{z_j}} \right]_{i,j=1}^{n}$$

is positive semidefinite and $\det P > 0$. *Fix* $z_{n+1} \in \mathbb{D} \backslash \{z_1, z_2, \ldots, z_n\}$. *Then*

$$W = \{ f(z_{n+1}) : f \in \mathscr{S}, f(z_k) = w_k, 1 \leqslant k \leqslant n \}$$

is a closed disk of positive radius in $\mathbb{D}$. *If f is a finite Blaschke product of degree n that satisfies*

$$f(z_k) = w_k, \qquad 1 \leqslant k \leqslant n,$$

then $f(z_{n+1}) \in \partial W$. *Conversely, if* $w_{n+1} \in \partial W$, *then there is a unique finite Blaschke product f of degree n such that*

$$f(z_k) = w_k, \qquad 1 \leqslant k \leqslant n+1.$$

Proof Recall from Lemma 2.1.6 that if W is a closed disk in $\mathbb{D}$ and $\tau \in \operatorname{Aut}(\mathbb{D})$, then $\tau(W)$ is also a disk in $\mathbb{D}$. Moreover, τ maps ∂W onto $\partial \tau(W)$ bijectively.

Our proof is by induction. However, even the base case $n = 1$ is rich enough to give a panoramic view of the whole process. As in the proof of Theorem 7.3.6, for fixed $z_1, w_1 \in \mathbb{D}$, there is an $f \in \mathscr{S}$ such that

$$f(z_1) = w_1 \tag{7.3.21}$$

if and only if

$$f = \tau_{w_1} \circ g \circ \tau_{z_1} \tag{7.3.22}$$

for some $g \in \mathscr{S}$ and $g(0) = 0$. Therefore we obtain an infinite number of solutions (7.3.22) to (7.3.21) as g runs over all the Schur class functions vanishing at the origin. This parameterization of solutions implies that for $z_2 \in \mathbb{D} \backslash \{z_1\}$

$$\begin{aligned} W &= \{f(z_2) : f \in \mathscr{S} \text{ satisfies } (7.3.21)\} \\ &= \{\tau_{w_1} \circ g \circ \tau_{z_1}(z_2) : g \in \mathscr{S}, g(0) = 0\} \\ &= \{\tau_{w_1}(g(\eta)) : g \in \mathscr{S}, g(0) = 0, \eta = \tau_{z_1}(z_2)\}. \end{aligned} \tag{7.3.23}$$

By the Schwarz Lemma (Lemma 1.1.1)

$$\{g(\eta) : g \in \mathscr{S}, g(0) = 0\} = D(0, |\eta|)^-.$$

Furthermore, if $\zeta \in \partial D(0, |\eta|)^-$ and $g \in \mathscr{S}$ with $g(0) = 0$ and $g(\eta) = \zeta$, then $|g(\eta)| = |\eta|$. Thus, $g(z) = \gamma z$ for some $\gamma \in \mathbb{T}$ by the Schwarz Lemma. In fact, each point on $\partial D(0, |\eta|)^-$ corresponds to a unique $\gamma \in \mathbb{T}$ via $g(z) = \gamma z$.

By the remarks at the beginning of the proof,

$$W = \tau_{w_1}(D(0, |\eta|)^-)$$

is a closed disk in $\mathbb{D}$ and

$$\partial W = \tau_{w_1}(\partial D(0, |\eta|)^-). \tag{7.3.24}$$

Suppose that f is a finite Blaschke product of degree one (an automorphism) with $f(z_1) = w_1$. Then by (7.3.22) $f = \tau_{w_1} \circ \rho_\gamma \circ \tau_{z_1}$ for some $\gamma \in \mathbb{T}$. By (7.3.23) and (7.3.24) we have

$$f(z_2) = \tau_{w_1}(\gamma\eta) \in \partial W.$$

Conversely if $w_2 \in \partial W$, then $w_2 = \tau_{w_1}(\gamma\eta)$ for some $\gamma \in \mathbb{T}$. If

$$f = \tau_{w_1} \circ \rho_\gamma \circ \tau_{z_1},$$

then f is a finite Blaschke product of degree one with $f(z_1) = w_1$ and

$$f(z_2) = \tau_{w_1}(\gamma \tau_{z_1}(z_2)) = \tau_{w_1}(\gamma\eta) = w_2.$$

This establishes the base case $n = 1$.

For the inductive step, suppose that $z_1, z_2, \ldots, z_n \in \mathbb{D}$ and $w_1, w_2, \ldots, w_n \in \mathbb{D}$ are given. Then $f \in \mathscr{S}$ is a solution to

$$f(z_j) = w_j, \quad 1 \leqslant j \leqslant n, \tag{7.3.25}$$

if and only if

$$f = \tau_{w_n} \circ (zg) \circ \tau_{z_n} \tag{7.3.26}$$

for some $g \in \mathscr{S}$ with

$$g(z'_j) = w'_j, \quad 1 \leqslant j \leqslant n-1, \tag{7.3.27}$$

where

$$z'_j = \tau_{z_n}(z_j) \quad \text{and} \quad w'_j = \frac{\tau_{z_n}(w_j)}{\tau_{w_n}(z_j)}.$$

If $z_{n+1} \in \mathbb{D}\backslash\{z_1, z_2, \ldots, z_n\}$, then

$$\begin{aligned} W &= \{f(z_{n+1}) : f \in \mathscr{S} \text{ satisfies (7.3.25)}\} \\ &= \{\tau_{w_n}(z'_{n+1} g(z'_{n+1})) : g \in \mathscr{S} \text{ satisfies (7.3.27)}\}, \end{aligned}$$

where $z'_{n+1} = \tau_{z_n}(z_{n+1}) \neq 0$. As $g \in \mathscr{S}$ runs through the solutions to (7.3.27), the inductive hypothesis says that

$$\{g(z'_{n+1}) : g \in \mathscr{S} \text{ satisfies (7.3.27)}\}$$

is a closed disk in $\mathbb{D}$ and each point on its boundary is of the form $g(z'_{n+1})$, in which g is the unique finite Blaschke product of degree $n-1$ satisfying (7.3.27). Thus, W is also a closed disk in $\mathbb{D}$ and, considering (7.3.26), each point of its boundary is of the form $f(z_{n+1})$, where f is the unique finite Blaschke product of degree n satisfying (7.3.25).

Conversely, if f is a Blaschke product of degree n with

$$f(z_j) = w_j, \quad 1 \leqslant j \leqslant n+1,$$

then, via (7.3.26), there is a finite Blaschke product g of degree $n-1$ for which

$$g(z'_j) = w'_j, \quad 1 \leqslant j \leqslant n-1,$$

and

$$g(z'_{n+1}) = w'_{n+1}.$$

Therefore,

$$w'_{n+1} \in \partial\{g(z'_{n+}) : g \text{ satisfies (7.3.27)}\},$$

which in turn implies that $w_{n+1} \in \partial W$. □

7.4 Boundary Interpolation: Cantor–Phelps Solution

Given distinct $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ and arbitrary $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{T}$, is there a finite Blaschke product B so that

$$B(\zeta_k) = \xi_k, \qquad 1 \leqslant k \leqslant n?$$

It turns out that this interpolation problem is always solvable. In this section, we provide an existence proof that involves the Cantor–Phelps theorem, a remarkable general result. In the next section, we give a more constructive approach.

Let S be a *semigroup*, that is, a set endowed with an associative binary operation and an identity element. Examples include $\mathbb{N}$ with the operation of multiplication or M_n with the operation of (matrix) multiplication. The semigroup that will be important here is $\mathbb{T}$ under multiplication. In fact, $\mathbb{T}$ is a group since each $z \in \mathbb{T}$ has the multiplicative inverse $\overline{z} \in \mathbb{T}$.

Let $\mathscr{F}$ be a collection of functions from a semigroup S into itself. If $f, g \in \mathscr{F}$, then both $f \circ g$ and fg are well-defined functions on S. More precisely, when writing $(fg)(\alpha) = f(\alpha)g(\alpha)$ we use the binary operation on S to calculate $f(\alpha)g(\alpha)$. In the following, we consider families $\mathscr{F}$ that are closed with respect to both operations; that is,

$$f, g \in \mathscr{F} \quad \Longrightarrow \quad f \circ g \in \mathscr{F} \quad \text{and} \quad fg \in \mathscr{F}.$$

Definition 7.4.1 Let $\mathscr{F}$ be a collection of functions from a semigroup S into itself. Then $\mathscr{F}$ is *n-transitive* if for any n distinct $\alpha_1, \alpha_2, \cdots, \alpha_n \in S$ and any arbitrary $\beta_1, \beta_2, \ldots, \beta_n \in S$, there is an $f \in \mathscr{F}$ such that

$$f(\alpha_i) = \beta_i, \qquad 1 \leqslant i \leqslant n.$$

We say that $\mathscr{F}$ is *transitive* if it is n-transitive for all $n \geqslant 1$. Although transitive families can be defined without the assumption that the common domain S forms a semigroup, for our purposes, we also require $\mathscr{F}$ to be closed under pointwise products.

Is there a family that is n-transitive but not $(n+1)$-transitive? For $n = 1$, the answer is yes: the family of all constant functions on a semigroup is 1-transitive but not 2-transitive. See Exercise 7.13 for an example of a family that is 2-transitive but not 3-transitive. The following result shows that under some mild conditions, any 3-transitive family is transitive.

Lemma 7.4.2 (Cantor–Phelps [14]) *Let S be a semigroup with identity element* 1 *and suppose that S contains an element δ such that $1, \delta, \delta^2$ are distinct elements of S. If $\mathscr{F}$ is a collection of functions from S into itself that is closed under composition, closed under pointwise multiplication, and* 3*-transitive, then $\mathscr{F}$ is transitive.*

Proof Fix $n \geqslant 4$. Given any n distinct elements $\alpha_1, \alpha_2, \ldots, \alpha_n \in S$ and arbitrary $\beta_1, \beta_2, \ldots, \beta_n \in S$, we need to show that there is an $f \in \mathscr{F}$ such that

$$f(\alpha_i) = \beta_i, \qquad 1 \leqslant i \leqslant n. \tag{7.4.3}$$

Since S has an identity element and $\mathscr{F}$ is closed under pointwise multiplication, it suffices to show that there are $f_1, f_2, \ldots, f_n \in \mathscr{F}$ such that

$$f_j(\alpha_i) = \begin{cases} \beta_j & \text{if } i = j, \\ 1 & \text{if } i \neq j. \end{cases}$$

Indeed, the function

$$f = f_1 f_2 \cdots f_n \tag{7.4.4}$$

so constructed will solve the interpolation problem (7.4.3). We now show the existence of f_1. The other cases are similar.

Suppose that $\mathscr{F}$ is $(n-1)$-transitive. Then there are $g, h \in \mathscr{F}$ such that $g(\alpha_1) = h(\alpha_1) = \delta$,

$$g(\alpha_i) = 1, \qquad 2 \leqslant i \leqslant n-1,$$

and

$$h(\alpha_i) = 1, \qquad 3 \leqslant i \leqslant n;$$

we have no control over $g(\alpha_n)$ and $h(\alpha_2)$. For simplicity, let

$$\gamma = g(\alpha_n) \quad \text{and} \quad \gamma' = h(\alpha_2).$$

Based on the values of γ and γ' there are three special cases.

(a) Suppose that $\gamma \neq \delta$. Since $\mathscr{F}$ is 3-transitive, there is a $k \in \mathscr{F}$ such that $k(\delta) = \beta_1$ and $k(1) = k(\gamma) = 1$. Let $f_1 = k \circ g$.
(b) Suppose that $\gamma' \neq \delta$. This is similar to the preceding case. We have a $k \in \mathscr{F}$ such that $k(\delta) = \beta_1$ and $k(1) = k(\gamma') = 1$. Let $f_1 = k \circ h$.
(c) Suppose that $\gamma = \gamma' = \delta$. In this situation, $gh \in \mathscr{F}$ maps α_1 to δ^2, both α_2 and α_n to δ, and the other arguments to 1. Hence, we pick up a $k \in \mathscr{F}$ such that $k(\delta^2) = \beta_1$ and $k(1) = k(\delta) = 1$. Now take $f_1 = k \circ (gh)$.

Thus, f_1 has the desired properties and hence $\mathscr{F}$ is n-transitive.

Since $\mathscr{F}$ is 3-transitive and since $\mathscr{F}$ is n-transitive whenever it is $(n-1)$-transitive, induction guarantees that $\mathscr{F}$ is transitive. □

As a consequence of Lemma 7.4.2, we obtain an existence result for boundary interpolation by finite Blaschke products.

Theorem 7.4.5 *The family of all finite Blaschke products, considered as functions on $\mathbb{T}$, is transitive.*

Proof The unit circle $\mathbb{T}$ is a group with identity 1. If $\delta = i$,then $1, \delta, \delta^2$ are distinct. Theorem 3.6.2 ensures that the family $\mathscr{F}$ of all finite Blaschke products is closed under pointwise multiplication and composition. By Lemma 7.4.2, it suffices to show that $\mathscr{F}$ is 3-transitive.

First suppose that $\alpha_1, \alpha_2 \in \mathbb{T}$ are distinct and $\beta_1, \beta_2 \in \mathbb{T}$ are arbitrary. By pre- and post-composing the desired interpolating function with appropriate rotations (which are finite Blaschke products of order 1), we may assume that

$$\alpha_2 = \overline{\alpha_1}, \qquad \beta_2 = \overline{\beta_1}, \qquad \operatorname{Im}\alpha_1 > 0, \quad \text{and} \quad \operatorname{Im}\beta_1 \geqslant 0. \tag{7.4.6}$$

There are two cases to consider.

(a) If $\beta_1 \neq \beta_2$, let

$$a = \frac{\beta_1 - \alpha_1}{1 - \beta_1\alpha_1}.$$

Then $\overline{a} = a$ and

$$1 - a^2 = \frac{4(\operatorname{Im}\alpha_1)(\operatorname{Im}\beta_1)}{|1 - \alpha_1\beta_1|^2} > 0$$

by (7.4.6). Thus, $a \in (-1, 1)$ and the Mobius transformation

$$f(z) = \frac{z - a}{1 - az},$$

satisfies $f(\alpha_1) = \beta_1$ and $f(\alpha_2) = \beta_2$.

(b) If $\beta_1 = \beta_2$, then (7.4.6) implies that $\beta_1 = \beta_2 = 1$. Although the constant function 1 solves the two-point interpolation problem $f(\alpha_1) = f(\alpha_2) = 1$, it is not ideal for our future applications. Use (a) to produce a disk automorphism g such that $g(\alpha_1) = 1$ and $g(\alpha_2) = -1$. Then $f = g^2$ is a finite Blaschke product of degree 2 such that $f(\alpha_1) = f(\alpha_2) = 1$. Since f is a Blaschke product of degree 2, Theorem 3.4.10 implies that $f(\alpha) \neq 1$ for any $\alpha \in \mathbb{T}\backslash\{\alpha_1, \alpha_2\}$.

Suppose that $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{T}$ are distinct and $\beta_1, \beta_2, \beta_3 \in \mathbb{T}$ are arbitrary. Mimicking the construction (7.4.4) in the proof of Lemma 7.4.2, we may assume that $\beta_1 = \beta_2 = 1$ and $\beta_3 \in \mathbb{T}$ is arbitrary. By (b) of the preceding discussion, there is a finite Blaschke product g such that $g(\alpha_1) = g(\alpha_2) = 1$ and $g(\alpha_3) \neq 1$. By (a), there is an $h \in \operatorname{Aut}(\mathbb{D})$ such that $h(1) = 1$ and $h(g(\alpha_3)) = \beta_3$. Then $f = h \circ g$ satisfies $f(\alpha_i) = \beta_i$ for $i = 1, 2, 3$. Thus, $\mathscr{F}$ is 3-transitive. □

We repeat the content of Theorem 7.4.5 in the more familiar language of interpolation below. This version is more appropriate in our context.

Theorem 7.4.7 *Let $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ be distinct and let $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{T}$ be arbitrary. Then there is a finite Blaschke product B such that*

$$B(\zeta_i) = \xi_i, \qquad 1 \leqslant i \leqslant n. \tag{7.4.8}$$

Another nonconstructive approach to the preceding theorem that also optimizes the degree of the interpolating Blaschke product is due to Jones and Ruscheweyh [82]. See also [78].

7.5 Boundary Interpolation: A Constructive Solution

Theorem 7.4.7, which concerns boundary interpolation by finite Blaschke products, deserves more attention. The proof that we gave, which depends upon the Cantor–Phelps lemma (Lemma 7.4.2), does not provide a transparent construction. We provide a more constructive approach in this section.

Let $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{T}$ be distinct and let $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{T}$ be arbitrary. We wish to produce a finite Blaschke product B that satisfies (7.4.8). Pick $\zeta \in \mathbb{T}\backslash\{\zeta_1, \zeta_2, \ldots, \zeta_n\}$. We apply the Möbius transformation

$$\varphi(z) = i\,\frac{\zeta + z}{\zeta - z}$$

to transfer our problem from $\mathbb{T}$ to $\mathbb{R}$. Observe how φ provides bijective mappings between $\mathbb{D}$ and $\mathbb{C}_+$, between $\mathbb{C}\backslash\mathbb{D}^-$ and $\mathbb{C}_-$, and between $\mathbb{T}$ and $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$.

One constructive approach to our interpolation problem originates in [138]. Exercise 7.5 outlines the construction of a rational function f so that $B = \varphi^{-1} \circ f \circ \varphi$ is a finite Blaschke product that solves (7.4.8). A closer look at the construction reveals that that the order of f, and hence the order of B, is at most $n^2 - n$. However, it is natural to wonder if we can do better.

There are n free parameters that determine a Blaschke product of order $n-1$; these are the $n-1$ zeros and a unimodular constant factor. Consequently, we expect that the boundary interpolation problem (7.4.8) of Theorem 7.4.7 has a solution that is a finite Blaschke product of order at most $n-1$. This is obtained using the methods introduced in [65], which we describe below.

As above, apply the conformal mapping φ to reduce the problem to an interpolation problem on $\mathbb{R}$. Given distinct $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ and arbitrary $\beta_1, \beta_2, \ldots, \beta_n \in \mathbb{R}$, we seek a rational function f of order at most $n-1$ such that

$$f(\alpha_i) = \beta_i, \qquad 1 \leqslant i \leqslant n;$$

moreover, f must map $\mathbb{C}_+$, $\mathbb{C}_-$, and $\widehat{\mathbb{R}}$ into themselves. The idea is to apply Corollary 7.2.6 with some of the α_k playing the role of the x_k and the rest the acting as a_ks. Hence the function f would be of order at most $n-1$.

Without loss of generality, we assume that

$$\alpha_1 < \alpha_2 < \cdots < \alpha_n.$$

If all of the β_k are the same, then the interpolation problem can be solved by a constant function. If not, there is a j such that

$$\beta_j \neq \beta_{j+1} = \beta_{j+2} = \cdots = \beta_n.$$

We partition $\alpha_1, \alpha_2, \ldots, \alpha_n$ into two sets E_1 and E_2 defined by

$$E_1 = \{\alpha_i : \beta_i = \beta_j\} \quad \text{and} \quad E_2 = \{\alpha_1, \alpha_2, \ldots, \alpha_n\} \backslash E_1;$$

observe that each set is nonempty and

$$|E_1| + |E_2| = n.$$

If the points of E_1 and E_2 are interlaced, then their elements can play the role of a_ks and x_ks, respectively, in Corollary 7.2.6. In this case, we are done. In general, there is no reason to believe that the elements of E_1 and E_2 interlace appropriately. If this occurs, then we must add extra points in order to interlace the elements of E_1 and E_2.

Suppose that E_1 is partitioned into m nonempty subsets so that the elements of each subset are adjacent, with respect to the natural ordering on $\mathbb{R}$, and so that two neighboring subsets are separated by at least one element of E_2. Denote the number of points in these subsets by $\ell_1, \ell_2, \ldots, \ell_m$ and observe that

$$\ell_1 + \cdots + \ell_m = |E_1|.$$

The kth subset consists of ℓ_k points

$$\alpha_{i_k} < \alpha_{i_k+1} < \cdots < \alpha_{i_k+\ell_k}$$

with $\alpha_{i_k+\ell_k+1} \in E_2$. If $i_1 > 1$, then $\alpha_{i_k-1} \in E_2$. Select $\ell_k - 1$ real numbers $t_1, t_2, \ldots, t_{\ell_k-1}$ such that

$$\alpha_{i_k} < t_1 < \alpha_{i_k+1} < t_2 < \cdots < t_{\ell_k-1} < \alpha_{i_k+\ell_k}$$

and let

$$\mathscr{E}_2 = E_2 \cup \Big(\bigcup_{k=1}^{m} \{t_1, t_2, \ldots, t_{\ell_k-1}\} \Big).$$

Then

$$|\mathscr{E}_2| = |E_2| + (\ell_1 - 1) + \cdots + (\ell_m - 1) = |E_2| + |E_1| - m = n - m.$$

Now consider E_1 and $\mathscr{E}_2$ together. The elements of E_1 and $\mathscr{E}_2$ are intended to play the role of a_k and x_k in Corollary 7.2.6. At present, we have sufficiently many candidates x_k so that no two candidates a_k are adjacent. However, some of the elements of E_1 may be adjacent to each other. Suppose that these determine m' line segments, none of which contain any elements of E_1. Denote the number of elements of these segments by $\ell'_1, \ell'_2, \ldots, \ell'_{m'}$ so that

$$\ell'_1 + \cdots + \ell'_{m'} = |\mathscr{E}_2| = n - m.$$

Since $\alpha_n \notin E_1$, we have

$$m' = \begin{cases} |E_1| & \text{if } \alpha_1 \in E_1, \\ |E_1| + 1 & \text{if } \alpha_1 \notin E_1. \end{cases}$$

As we did above for E_2, enlarge E_1 to obtain an $\mathscr{E}_1$ so that the elements of $\mathscr{E}_1$ and $\mathscr{E}_2$ are interlaced. Then

$$|\mathscr{E}_1| = |E_1| + (\ell'_1 - 1) + \cdots + (\ell'_{m'} - 1) = |E_1| - m' + n - m.$$

There are two possibilities.

(a) If $|E_1| = m'$, then we have two interlaced sets $\mathscr{E}_1$ and $\mathscr{E}_2$, each with $n - m$ elements, and so that the first element of $\mathscr{E}_1 \cup \mathscr{E}_2$ belongs to $\mathscr{E}_1$.
(b) If $|E_1| = m' - 1$, then the two sets $\mathscr{E}_1$ and $\mathscr{E}_2$ are interlaced, but the first element of $\mathscr{E}_1 \cup \mathscr{E}_2$ belongs to $\mathscr{E}_2$. In this case, one last modification is needed. Pick an element that is smaller than any point in $\mathscr{E}_1 \cup \mathscr{E}_2$ and add it to $\mathscr{E}_1$. For simplicity, we label this new set $\mathscr{E}_1$. Then we have two interlaced sets $\mathscr{E}_1$ and $\mathscr{E}_2$, each having $n - m$ elements, and the first element of $\mathscr{E}_1 \cup \mathscr{E}_2$ belongs to $\mathscr{E}_1$.

All the elements of $\mathscr{E}_2$ should be mapped to the y_js. Some elements of $\mathscr{E}_1$ are the α_k of Corollary 7.2.6 and we know where they must be mapped to. For the image of the rest, pick arbitrary real numbers $\mathbb{R}$. Then Corollary 7.2.6 yields the existence of a rational function f whose order is $n - m$ that performs the desired interpolation. If $m > 1$, we can add $m - 1$ extra appropriate points to both $\mathscr{E}_1$ and $\mathscr{E}_2$ so that we obtain a solution of order $n - 1$.

As the construction above shows, if $\{\beta_1, \beta_2, \ldots, \beta_n\}$ is not a singleton, then it is possible to find a Blaschke product B of order $n - 1$ that performs the boundary interpolation (7.4.8). This hypothesis cannot be relaxed. For example, if $\alpha_1 = 1$ and $\alpha_2 = -1$, then there is no Blaschke product B of degree one such that $\beta = B(1) = B(-1)$; this is a consequence of Theorem 3.4.10. As a matter of fact, if

$$B(z) = \gamma \frac{z_0 - z}{1 - \overline{z_0}\, z}, \quad \gamma \in \mathbb{T},\ z_0 \in \mathbb{D},$$

then the assumption $B(1) = B(-1)$ implies $|z_0|^2 = 1$, which is absurd. Hence, the only solution for the interpolation problem is the unimodular constant function $B \equiv \beta$.

7.6 Exercises

7.1 Show that the *Vandermonde matrix*

$$V(z_1, z_2, \ldots, z_n) = \begin{bmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{n-1} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & z_n & z_n^2 & \cdots & z_n^{n-1} \end{bmatrix}$$

is invertible if and only if $z_1, z_2, \ldots, z_n$ are distinct. Use this to prove the Lagrange interpolation theorem (Theorem 7.1.1).

7.2 Let $A \in M_n$ with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ and corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r \in \mathbb{C}^n$. Prove that $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r$ are linearly independent.
Hint: Suppose that $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_r\mathbf{x}_r = \mathbf{0}$. Use Lagrange interpolating polynomials to show that all of the c_j vanish.

7.3 Suppose that $A \in M_n$ has distinct eigenvalues. Let $\{A\}'$ denote the *commutant* of A, the set of all matrices that commute with A. Prove that $\{A\}' = \{p(A) : p \text{ is a polynomial}\}$ and $\dim\{A\}' = n$.
Hint: Use the Lagrange interpolation theorem and the fact that A is diagonalizable.

7.4 Suppose that $A_1, A_2, \ldots, A_r \in M_n$ are normal matrices: $A_j^* A_j = A_j A_j^*$ for each j. Prove that there is a polynomial p so that $A_i^* = p(A_i)$ for $i = 1, 2, \ldots, k$.
Hint: Use the Lagrange interpolation theorem.

7.5 (Younis [138]) Let $x_1, x_2, \ldots, x_n$ be a finite sequence of distinct real numbers, and let $y_1, y_2, \ldots, y_n$ be real numbers. Let

$$p_k(z) = y_k - \sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{1}{x_j - x_k} + \sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{1}{x_j - z}$$

and

$$q_k(z) = \frac{-y_k^2}{z}$$

for $1 \leqslant k \leqslant n$. Define

$$f = (q_1 \circ p_1) + (q_2 \circ p_2) + \cdots + (q_n \circ p_n).$$

Show that f is a rational function that satisfies the following.

(a) $f(\mathbb{C}_+) \subseteq \mathbb{C}_+$.
(b) $f(\widehat{\mathbb{R}}) \subseteq \widehat{\mathbb{R}}$. In particular, f only has real poles and real zeros.
(c) $f(x_k) = y_k$ for all $1 \leqslant k \leqslant n$.

Give an upper bound for the order of f.

Remark This result is a weaker version of Corollary 7.2.6. However, its proof is simpler.

7.6 Show that $A \in M_n$ is positive semidefinite if and only if $A = A^*$ and all of its eigenvalues are nonnegative.

7.7 Suppose that $A \in M_n$ and positive semidefinite and S is $m \times n$. (a) Show that SAS^* is positive semidefinite. (b) If $S \in M_n$ is invertible, show that A is positive semidefinite if and only if SAS^* is positive semidefinite.

7.8 Verify (7.3.11) and (7.3.12).

7.9 Suppose that $z_1, z_2, z_3 \in \mathbb{D}$ are distinct and $w_1, w_2, w_3 \in \mathbb{D}$ are arbitrary. If there are finite Blaschke products f_1, f_2, f_3 so that $f_j(z_i) = w_i$ for $i \neq j$, does there exist a finite Blaschke product f so that $f(z_i) = w_i$ for $i = 1, 2, 3$?

7.10 Let $z_1, z_2, \ldots, z_n$ be a set of distinct points in $\mathbb{C}$. For each $i = 1, 2, \ldots, n$, let w_{ij}, for $j = 0, 1, \ldots, J(i)$, be an arbitrary set of complex numbers. Find the unique polynomial of degree $n - 1 + J(1) + \cdots + J(n)$ such that

$$f^{(j)}(z_i) = w_{ij}, \quad 1 \leqslant i \leqslant n, \quad 0 \leqslant j \leqslant J(i).$$

This is known as *Hermite interpolation* [87].
Hint: Consider linear combination of polynomials

$$(z - z_1)^{k_1}(z - z_2)^{k_2} \cdots (z - z_n)^{k_n}.$$

Remark The case studied in this chapter (Lagrange interpolation) corresponds to $J(i) = 0$ for all i.

7.11 Let $z_1, z_2, z_3 \in \mathbb{C}$ be distinct and let $w_1, w_2, w_3 \in \mathbb{C}$ be arbitrary. Show that there is a unique Möbius transformation f such that $f(z_k) = w_k$ for $k = 1, 2, 3$.
Hint: Consider

$$\frac{(w - w_1)(w_3 - w_2)}{(w - w_2)(w_3 - w_1)} = \frac{(z - z_1)(z_3 - z_2)}{(z - z_2)(z_3 - z_1)}.$$

7.12 Use the Pick interpolation theorem (Theorem 7.3.6) to give a new proof of Carathéodory's theorem (Theorem 4.1.1).

Hint: Pick $z_1, z_2, \ldots, z_n$ to be the vertices of a regular n-gon that is inscribed in the circle $|z| = \varepsilon$ and find a finite Blaschke product B such that $B(z_k) = f(z_k)$ for all k. Then $f - B = Cg$, where $g \in \mathscr{S}$ and C is a finite Blaschke product with zeros at z_k. Hence $|f - B| \leqslant |C|$. Fix a compact set K and let $\epsilon \to 0$ and $n \to \infty$.

7.13 Let $\mathscr{F}$ be the family of all linear maps on $\mathbb{R}$ with the operation of composition. Show that $\mathscr{F}$ is 2-transitive, but not 3-transitive.

7.14 The goal of this exercise is to show that in Lemma 7.4.2 the requirement that $1, \beta, \beta^2$ are distinct cannot be relaxed. Let $G = \{1, a, b, c\}$ be the Klein four-group, whose multiplication table is

$\times$	1	a	b	c
1	1	a	b	c
a	a	1	c	b
b	b	c	1	a
c	c	b	a	1

and let $\mathscr{F}$ be the family of all functions $f : G \to G$ such that

$$f(1) \times f(a) \times f(b) \times f(c) = 1.$$

Observe that $\beta^2 = 1$ for each $\beta \in G$. Prove the following.

(i) $\mathscr{F}$ is closed under pointwise multiplication.
(ii) $\mathscr{F}$ is closed under composition.
(iii) $\mathscr{F}$ is 3-transitive.
(iv) $\mathscr{F}$ is not 4-transitive.

Hint: Observe that $\mathscr{F}$ contains precisely the following functions:

(a) the constant functions;
(b) the bijective functions;
(c) those functions whose range contains 2 elements of G, and each element in the range is the image of two elements in the domain.

7.15 Let $z_1, z_2, \ldots, z_n$ and $\zeta_1, \zeta_2, \ldots, \zeta_n$ be elements of $\mathbb{T}$ that satisfy

$$0 \leqslant \arg \zeta_1 < \arg z_1 < \arg \zeta_2 < \arg z_2 < \cdots < \arg \zeta_n < \arg z_n < 2\pi,$$

and let $w_0, w_1, \ldots, w_n \in \mathbb{T}$ be such that $w_0 \neq w_k$ for $1 \leqslant k \leqslant n$. Show that there is a finite Blaschke product of degree n such that

$$B(z_k) = w_k, \qquad 1 \leqslant k \leqslant n,$$

and

$$B(\zeta_k) = w_0, \qquad 1 \leqslant k \leqslant n.$$

Hint: Use Corollary 7.2.6.

Chapter 8
The Bohr Radius

The Bohr radius was examined over a century ago by H. Bohr [9] and it is still a source of inspiration and further studies. We take up this subject for two reasons. First, the solution to certain extremal problems involves either disk automorphisms or a finite Blaschke product of order two. Second, there are several questions in this area that are not yet settled and it is conjectured that the solution to these problems should involve a finite Blaschke product.

Recall from (1.0.2) the *Schur class* $\mathscr{S}$ of all analytic functions $f : \mathbb{D} \to \mathbb{D}^-$. For $f \in \mathscr{S}$ with Taylor series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

at the origin, one can ask about the possible values of $r \in [0, 1)$ such that

$$\sum_{n=0}^{\infty} |a_n| r^n \leqslant 1, \qquad f \in \mathscr{S}. \tag{8.0.1}$$

This inequality holds when $r = 0$ since

$$|a_0| = |f(0)| \leqslant 1, \qquad f \in \mathscr{S}.$$

H. Bohr [9] observed that (8.0.1) is true when $r \in [0, \frac{1}{6}]$. This work was expanded further by M. Riesz, Schur, and Wiener [32] who independently showed that (8.0.1) holds when $r \in [0, \frac{1}{3}]$ and that $\frac{1}{3}$, now called the *Bohr radius* $\mathfrak{B}_0$, is the best possible constant.

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_8

8.1 The Classical Bohr Radius $\mathfrak{B}_0$

In this section we present four equivalent definitions of the Bohr radius $\mathfrak{B}_0$. Each point of view has its own merits and is important for further generalizations. In subsequent sections, we will provide several proofs of the fact that $\mathfrak{B}_0 = \frac{1}{3}$.

For $f \in \mathscr{S}$, let

$$m(f, r) := \sum_{n=0}^{\infty} |a_n| r^n, \qquad r \in [0, 1). \tag{8.1.1}$$

Observe that for each $f \in \mathscr{S}$ the function $r \mapsto m(f, r)$ is increasing on $[0, 1)$ with

$$m(f, 0) = |a_0| \leqslant 1.$$

It is important to note that for some $f \in \mathscr{S}$, the function $m(f, r)$ assumes values larger than 1. To see this, let b be the disk automorphism

$$b(z) = \frac{a - z}{1 - az}, \tag{8.1.2}$$

in which a is a free parameter in $(0, 1)$. Then $b \in \mathscr{S}$ and a geometric series computation confirms that

$$b(z) = a + (a^2 - 1) \sum_{n=1}^{\infty} a^{n-1} z^n. \tag{8.1.3}$$

Thus,

$$m(b, r) = \frac{a + (1 - 2a^2) r}{1 - ar}, \qquad r \in [0, 1), \tag{8.1.4}$$

and hence $m(b, r) > 1$ whenever $r > (1 + 2a)^{-1}$. This allows us to make the following definition.

Definition 8.1.5 The *Bohr radius* is the unique value $\mathfrak{B}_0 \in [0, 1]$ that satisfies the following.

(a) For all $f = \sum_{n=0}^{\infty} a_n z^n \in \mathscr{S}$,

$$\sum_{n=0}^{\infty} |a_n| r^n \leqslant 1, \qquad r \in [0, \mathfrak{B}_0].$$

(b) For each $r \in (\mathfrak{B}_0, 1)$, there is an $f = \sum_{n=0}^{\infty} a_n z^n \in \mathscr{S}$ such that

$$\sum_{n=0}^{\infty} |a_n| r^n > 1.$$

Although we have claimed, in our introductory remarks, that $\mathfrak{B}_0 = \frac{1}{3}$, at this point we can only conclude that $\mathfrak{B}_0$ exists and belongs to $[0, 1)$. This will be remedied shortly.

Definition 8.1.6 Suppose that $f = \sum_{n=0}^{\infty} a_n z^n \in \mathscr{S}$. If $\sum_{n=0}^{\infty} |a_n| > 1$, define $c = c(f)$ to be the unique value in $[0, 1)$ such that

$$m(f, c) = 1; \tag{8.1.7}$$

otherwise, let $c(f) = 1$. For example, (8.1.4) shows that $m(b, c(b)) = 1$ when

$$c(b) = \frac{1}{1 + 2a}. \tag{8.1.8}$$

Also define

$$\mathfrak{M}(\mathscr{S}, r) := \sup_{f \in \mathscr{S}} m(f, r), \qquad r \in [0, 1), \tag{8.1.9}$$

and observe that $\mathfrak{M}(\mathscr{S}, 0) = 1$.

Since $\mathfrak{M}(\mathscr{S}, r)$ is an increasing function of r, we conclude that $\mathfrak{M}(\mathscr{S}, r) \geqslant 1$ for all r. Although the notation $\mathscr{S}$ in $\mathfrak{M}(\mathscr{S}, r)$ might appear redundant, later on we will define a similar quantity for sets of functions other than the Schur class $\mathscr{S}$. We leave it to the reader as an exercise (see Exercise 8.1) to prove the following.

Proposition 8.1.10 *The Bohr radius $\mathfrak{B}_0$ satisfies the following.*

(a) $\mathfrak{B}_0 = \inf\{c(f) : f \in \mathscr{S}\}$.
(b) $\mathfrak{B}_0$ *is the value for which* $\mathfrak{M}(\mathscr{S}, r) = 1$ *for* $r \in [0, \mathfrak{B}_0]$ *and* $\mathfrak{M}(\mathscr{S}, r) > 1$ *for* $r \in (\mathfrak{B}_0, 1)$.
(c) $\mathfrak{B}_0$ *is the largest* $r \in [0, 1)$ *for which* $\mathfrak{M}(\mathscr{S}, r) = 1$.

If f is an arbitrary bounded analytic function on $\mathbb{D}$, not necessarily in the Schur class, then

$$m(f, r) \leqslant \mathfrak{M}(\mathscr{S}, r)\|f\|_\infty, \qquad r \in [0, 1). \tag{8.1.11}$$

Therefore, computing or estimating $\mathfrak{M}(\mathscr{S}, r)$ is of relevance for understanding the rate of growth of bounded analytic functions. This is not an easy task. In fact, a precise formula for $\mathfrak{M}(\mathscr{S}, r)$ for all values of $r \in [0, 1)$ is unknown. Some elementary upper and lower bounds for $\mathfrak{M}(\mathscr{S}, r)$ are provided below. We also provide a formula for $\mathfrak{M}(\mathscr{S}, r)$ for certain values of r.

Before getting into the proof of the following lemma, the reader may wish to review Fatou's theorem concerning radial boundary values of bounded analytic functions (Theorem A.3.1 in Appendix A.3) along with some basic facts about Hardy spaces; see (A.4.1) in Appendix A.4.

Lemma 8.1.12 (Upper Bound)

$$\mathfrak{M}(\mathscr{S}, r) \leqslant \frac{1}{\sqrt{1-r^2}}, \qquad r \in [0, 1).$$

Proof Let $f = \sum_{n=0}^{\infty} a_n z^n \in \mathscr{S}$. By the Cauchy–Schwarz inequality,

$$\begin{aligned} m(f, r) &= \sum_{n=0}^{\infty} |a_n| r^n \\ &\leqslant \left(\sum_{n=0}^{\infty} |a_n|^2\right)^{\frac{1}{2}} \left(\sum_{n=0}^{\infty} r^{2n}\right)^{\frac{1}{2}} \\ &= \left(\int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi}\right)^{\frac{1}{2}} \frac{1}{\sqrt{1-r^2}} \qquad \text{(by (A.4.1))} \\ &\leqslant \|f\|_\infty \cdot \frac{1}{\sqrt{1-r^2}} \qquad \text{(by Theorem A.3.1)} \\ &\leqslant \frac{1}{\sqrt{1-r^2}}. \end{aligned}$$

Now take the supremum with respect to $f \in \mathscr{S}$ and obtain the desired result. □

Lemma 8.1.13 (Lower Bound)

$$\mathfrak{M}(\mathscr{S}, r) \geqslant \begin{cases} 1 & \text{if } 0 \leqslant r \leqslant \frac{1}{3}, \\ \dfrac{3 - \sqrt{8(1-r^2)}}{r} & \text{if } \frac{1}{3} \leqslant r < 1. \end{cases}$$

Proof The constant function $f \equiv 1$ belongs to $\mathscr{S}$ and so

$$m(\mathscr{S}, r) \geqslant 1, \qquad r \in [0, 1).$$

We now obtain a better estimate for $r \geqslant \frac{1}{3}$ by using the automorphism b from (8.1.2). Then (8.1.4) provides

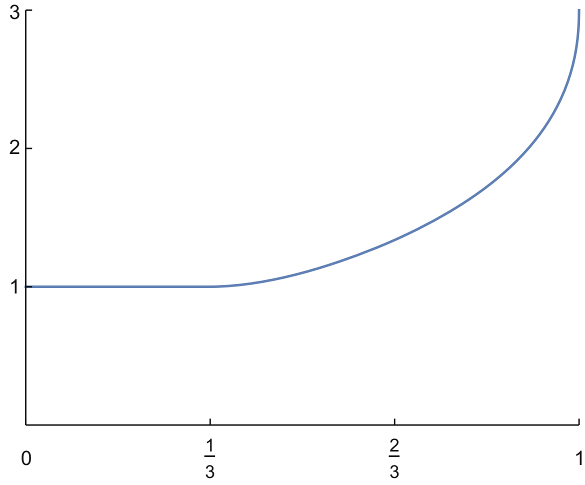


Fig. 8.1 The lower bound in Lemma 8.1.13

$$\mathfrak{M}(\mathscr{S}, r) \geqslant \frac{a + (1 - 2a^2)r}{1 - ar}, \qquad r \in [0, 1). \tag{8.1.14}$$

For each fixed r, we maximize the right-hand side in the preceding inequality. If $r \in [0, \frac{1}{3}]$, then (8.1.14) yields $\mathfrak{M}(\mathscr{S}, r) \geqslant 1$, which is already known. When $r \geqslant \frac{1}{3}$, the optimal value of

$$\frac{a + (1 - 2a^2)r}{1 - ar} \tag{8.1.15}$$

occurs when

$$a = \frac{2 - \sqrt{2(1 - r^2)}}{2r}.$$

Substituting a into (8.1.14) yields the required lower bound; see Fig. 8.1. □

8.2 Computing $\mathfrak{B}_0$

We provide two proofs of the fact $\mathfrak{B}_0 = \frac{1}{3}$. Our first proof is from [110] and relies on the following lemma that handles a special case.

Lemma 8.2.1 (Paulsen–Popescu–Singh) *Let*

$$f = \sum_{n=0}^{\infty} a_n z^n$$

be analytic on $\mathbb{D}$*. If* $\operatorname{Re} f \leqslant 1$ *and* $f(0) \geqslant 0$*, then* $m(f, r) \leqslant 1$ *whenever* $r \in [0, \frac{1}{3}]$.

Proof For fixed $r \in [0, 1)$ and $n \geqslant 1$,

$$\begin{aligned}
&\frac{1}{\pi}\int_0^{2\pi} \operatorname{Re}\Big(1 - f(re^{i\theta})\Big)\, e^{-in\theta} d\theta \\
&= \frac{1}{\pi}\int_0^{2\pi} \left(1 - \frac{1}{2} f(re^{i\theta}) - \frac{1}{2}\overline{f(re^{i\theta})}\right) e^{-in\theta} d\theta \\
&= \frac{1}{\pi}\int_0^{2\pi} \left(1 - \frac{1}{2}\sum_{k=0}^{\infty} a_k r^k e^{ik\theta} - \frac{1}{2}\sum_{k=0}^{\infty} \overline{a_k} r^k e^{-ik\theta}\right) e^{-in\theta} d\theta \\
&= \frac{1}{\pi}\int_0^{2\pi} e^{-in\theta} d\theta - \sum_{k=0}^{\infty} a_k r^k \int_0^{2\pi} e^{i(k-n)\theta}\frac{d\theta}{2\pi} - \sum_{k=0}^{\infty} \overline{a_k} r^k \int_0^{2\pi} e^{i(-k-n)\theta}\frac{d\theta}{2\pi} \\
&= -a_n r^n
\end{aligned}$$

by (A.4.4). Since $\operatorname{Re} f \leqslant 1$ and $a_0 = f(0) > 0$,

$$\begin{aligned}
|a_n| r^n &\leqslant \frac{1}{\pi}\int_0^{2\pi} |\operatorname{Re}(1 - f(re^{i\theta}))| d\theta \\
&= \frac{1}{\pi}\int_0^{2\pi} (1 - \operatorname{Re} f(re^{i\theta})) d\theta \\
&= 2(1 - a_0) \qquad (8.2.2)
\end{aligned}$$

by the mean value theorem for harmonic functions, or a power series computation based upon (A.4.4). Let $r \to 1^-$ to obtain the important estimate

$$|a_n| \leqslant 2(1 - a_0), \qquad n \geqslant 1. \qquad (8.2.3)$$

This can be strengthened significantly; see Lemma 8.2.8 below. If $r \in [0, \frac{1}{3}]$, then

$$m(f, r) = \sum_{n=0}^{\infty} |a_n| r^n \leqslant a_0 + 2(1 - a_0)\sum_{n=1}^{\infty}\frac{1}{3^n} = 1. \quad \square$$

We continue to follow [110] and apply Lemma 8.2.1 to obtain the following.

Theorem 8.2.4 $\mathfrak{B}_0 = \frac{1}{3}$.

Proof If $f \in \mathscr{S}$, then $\operatorname{Re} f \leqslant 1$ and a suitable unimodular constant multiple of f is nonnegative at the origin. Then (8.2.3) implies $m(f, r) \leqslant 1$ for $r \in [0, \frac{1}{3}]$ as above. Taking the supremum over $f \in \mathscr{S}$ yields

$$\mathfrak{M}(\mathscr{S}, r) \leqslant 1, \qquad r \in [0, \tfrac{1}{3}].$$

Thus, $\mathfrak{B}_0 \geqslant \frac{1}{3}$. Now let b denote the automorphism (8.1.2) and use (8.1.8) to get

$$\mathfrak{B}_0 = \inf_{f \in \mathscr{S}} c(f) \leqslant \inf_{0 \leqslant a < 1} \frac{1}{1 + 2a} = \frac{1}{3}. \tag{8.2.5}$$

This completes the proof. □

The proof above provides disk automorphisms b for which $c(b)$ is arbitrarily close to $\frac{1}{3}$. It is natural to wonder whether there is an $f \in \mathscr{S}$ such that $c(f) = \frac{1}{3}$. The following corollary shows that no such extremal function exists.

Corollary 8.2.6 *There is no $f \in \mathscr{S}$ for which $c(f) = \frac{1}{3}$.*

Proof Suppose toward a contradiction that $f = \sum_{n=0}^{\infty} a_n z^n \in \mathscr{S}$ and $c(f) = \frac{1}{3}$. Then (8.2.3) yields

$$1 = \sum_{n=0}^{\infty} |a_n| (\tfrac{1}{3})^n \leqslant |a_0| + 2(1 - |a_0|) \sum_{n=1}^{\infty} \frac{1}{3^n} = 1$$

and hence

$$|a_n| = 2(1 - |a_0|), \qquad n \geqslant 1. \tag{8.2.7}$$

However, $f \in \mathscr{S}$ and hence Parseval's formula (A.4.3) ensures that

$$\sum_{n=0}^{\infty} |a_n|^2 = \int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi} \leqslant 1.$$

In particular, $a_n \to 0$, which along with (8.2.7), shows that $|a_0| = 1$ and $a_n = 0$ for all $n \geqslant 1$. In other words, f is a constant function that satisfies $c(f) = 1$, a contradiction. □

We now provide another proof of Theorem 8.2.4, also from [110]. This proof employs Wiener's inequality (Lemma 8.2.8), which is a strengthened version of (8.2.3). The advantage of this second method is that it permits us to compute some of the other Bohr coefficients; see Sect. 8.3.

Lemma 8.2.8 (Wiener's Inequality) *If*

$$f=\sum_{n=0}^{\infty}a_nz^n\in\mathscr{S},$$

then

$$|a_n|\leqslant 1-|a_0|^2,\qquad n\geqslant 1.$$

Proof Since $f\in\mathscr{S}$, we may consider the bounded linear operator on the Hardy space H^2 (see Appendix A.4) given by

$$T_f:H^2\to H^2,\qquad T_fg=fg;$$

this is a special example of a *Toeplitz operator* (see Appendix A.7). In fact, the operator norm $\|T_f\|$ of T_f satisfies $\|T_f\|=\|f\|_\infty$ (Theorem A.7.3). We can identify T_f with its matrix representation with respect to the orthonormal basis $1,z,z^2,\ldots$ of H^2, that is,

$$T_f=\begin{bmatrix} a_0 & 0 & 0 & 0 & \cdots\\ a_1 & a_0 & 0 & 0 & \cdots\\ a_2 & a_1 & a_0 & 0 & \cdots\\ a_3 & a_2 & a_1 & a_0 & \cdots\\ \vdots & \vdots & \vdots & \vdots & \ddots\end{bmatrix}; \tag{8.2.9}$$

see Exercise 8.4. Let $\mathscr{V}=\operatorname{span}\{1,z^n\}$ and let T be the compression of T_f to $\mathscr{V}$; that is $T=P_{\mathscr{V}}T_f|_{\mathscr{V}}$, in which $P_{\mathscr{V}}$ is the orthogonal projection of H^2 onto $\mathscr{V}$. Then (8.2.9) reveals that

$$T=\begin{bmatrix} a_0 & 0\\ a_n & a_0\end{bmatrix},$$

where we have identified T with its matrix representation with respect to the orthonormal basis $\{1,z^n\}$ of $\mathscr{V}$. Since $\|T\|\leqslant\|T_f\|\leqslant\|f\|_\infty\leqslant 1$, it follows that T is a contraction. This occurs precisely when

$$I-T^*T=\begin{bmatrix} 1-|a_0|^2-|a_n|^2 & -a_0\overline{a}_n\\ -\overline{a}_0a_n & 1-|a_0|^2\end{bmatrix}$$

is positive semidefinite; see (A.6.5). A computation confirms that

$$0\leqslant\det(I-T^*T)=(1-|a_0|^2)^2-|a_n|^2,$$

which implies the desired result. □

Proof (Second Proof of Theorem 8.2.4) Using disk automorphisms, we proved in (8.2.5) that $\mathfrak{B}_0 \leqslant \frac{1}{3}$. Hence it suffices to establish the reverse inequality. If $f = \sum_{n=0}^{\infty} a_n z^n \in \mathscr{S}$, then Lemma 8.2.8 yields

$$\begin{aligned} m(f, r) &= \sum_{n=0}^{\infty} |a_n| r^n \\ &\leqslant |a_0| + (1 - |a_0|^2) \sum_{n=1}^{\infty} r^n \\ &= |a_0| + (1 - |a_0|^2) \frac{r}{1 - r} \end{aligned}$$

whenever $r \in [0, 1)$. If $r = \frac{1}{3}$, then

$$m(f, \tfrac{1}{3}) \leqslant |a_0| + \tfrac{1}{2}(1 - |a_0|^2),$$

in which $|a_0| \leqslant 1$. Since the maximum of the function

$$x \mapsto x + \frac{1}{2}(1 - x^2)$$

on $[0, 1]$ is 1, we conclude that $m(f, \frac{1}{3}) \leqslant 1$. Take the supremum over $f \in \mathscr{S}$ to obtain $\mathfrak{M}(\mathscr{S}, \frac{1}{3}) \leqslant 1$. □

8.3 The Generalized Bohr Radius $\mathfrak{B}_k$

We follow the path laid out in Sect. 8.1 and generalize the concept of the Bohr radius, where the Schur class $\mathscr{S}$ is replaced by

$$z^k \mathscr{S} = \{z^k f : f \in \mathscr{S}\};$$

that is,

$$z^k \mathscr{S} = \{f \in \mathscr{S} : f(0) = f'(0) = \cdots = f^{(k-1)}(0) = 0\}. \tag{8.3.1}$$

Definition 8.3.2 Fix any integer $k \geqslant 0$. The *Bohr radius* of order k, denoted by $\mathfrak{B}_k$, is the constant $\mathfrak{B}_k \in [0, 1]$ that satisfies the following.

(a) For each $f = \sum_{n=0}^{\infty} a_n z^{n+k} \in z^k \mathscr{S}$,

$$\sum_{n=0}^{\infty} |a_n| r^{n+k} \leqslant 1, \qquad r \in [0, \mathfrak{B}_k].$$

(b) For each $r \in (\mathfrak{B}_k, 1)$, there is an $f = \sum_{n=0}^{\infty} a_n z^{n+k} \in z^k \mathscr{S}$ such that

$$\sum_{n=0}^{\infty} |a_n| r^{n+k} > 1.$$

Implicit in the preceding definition is the existence and uniqueness of the $\mathfrak{B}_k$. An approach analogous to that used in the discussion prior to Definition 8.1.5 justifies this apparent oversight; we leave the details to the reader. In terms of the quantity $c(f)$ from (8.1.7), we may write

$$\mathfrak{B}_k = \inf_{f \in \mathscr{S}} c(z^k f).$$

In terms of the quantity $\mathfrak{M}(\mathscr{S}, r)$ from (8.1.9), for $k \geqslant 1$ it follows that $\mathfrak{B}_k$ is the unique solution to

$$r^k \mathfrak{M}(\mathscr{S}, r) = 1. \tag{8.3.3}$$

Indeed, for $k \geqslant 1$ the function $r \mapsto r^k \mathfrak{M}(\mathscr{S}, r)$ is strictly increasing and hence $r^k \mathfrak{M}(\mathscr{S}, r) = 1$ has a unique solution in $(0, 1)$. Observe that Definition 8.3.2 is equivalent to the classical one when $k = 0$. For $k = 0$, we have $\mathfrak{M}(\mathscr{S}, r) \equiv 1$ on the interval $[0, \mathfrak{B}_0]$ and that is why in Proposition 8.1.10 we insisted that $\mathfrak{B}_0$ is the *largest* solution to $\mathfrak{M}(\mathscr{S}, r) = 1$.

From (8.3.1) we may also say that

$$\mathfrak{B}_k = \inf\{c(f) : f \in \mathscr{S},\ f(0) = f'(0) = \cdots = f^{(k-1)}(0) = 0\}.$$

Since $r^{k+1} < r^k$ on $(0, 1)$ and $\mathfrak{M}(\mathscr{S}, r)$ is increasing and satisfies $\mathfrak{M}(\mathscr{S}, r) > 1$ for $r > \mathfrak{B}_0$, it follows that $\mathfrak{B}_k < \mathfrak{B}_{k+1}$. Since $\mathfrak{M}(\mathscr{S}, r) \to \infty$ as $r \to 1^-$ by (8.6.17), we also see that $\mathfrak{B}_k \to 1$ as $k \to \infty$.

We now compute the Bohr coefficient $\mathfrak{B}_1$. The estimate

$$0.6 < \mathfrak{B}_1 < 0.7071$$

was first obtained by Ricci [117]. Fields Medalist Enrico Bombieri [10] gave an explicit formula for $\mathfrak{M}(\mathscr{S}, r)$ when $r \in [\frac{1}{3}, \frac{1}{\sqrt{2}}]$. This result is Theorem 8.6.15 and in particular it implies that $\mathfrak{B}_1 = \frac{1}{\sqrt{2}}$. We follow Paulsen–Popescu–Singh [110] and use the upper and lower bounds provided earlier to evaluate $\mathfrak{B}_1$.

Theorem 8.3.4 $\mathfrak{B}_1 = \frac{1}{\sqrt{2}}$.

Proof By Lemma 8.1.12,

$$\mathfrak{M}(\mathscr{S}, \tfrac{1}{\sqrt{2}}) \leqslant \sqrt{2},$$

and, by Lemma 8.1.13,

$$\mathfrak{M}(\mathscr{S}, \tfrac{1}{\sqrt{2}}) \geqslant \sqrt{2}.$$

Therefore

$$\tfrac{1}{\sqrt{2}}\mathfrak{M}(\mathscr{S}, \tfrac{1}{\sqrt{2}}) = 1.$$

This identity, together with (8.3.3), implies that $\mathfrak{B}_1 = \frac{1}{\sqrt{2}}$. □

8.4 A Localized Bohr Radius

Let us introduce another generalization of the Bohr radius, in which the Schur class $\mathscr{S}$ is replaced by

$$\mathscr{S}_\lambda := \{f \in \mathscr{S} : f(0) = \lambda\}, \qquad \lambda \in [0, 1).$$

Observe that each member of $\mathscr{S}_\lambda$ has a Taylor series representation of the form

$$f(z) = \lambda + a_1 z + a_2 z^2 + \cdots .$$

Definition 8.4.1 $\mathfrak{B}_0(\lambda)$ is the unique number in [0, 1] that satisfies the following.

(a) For all $f = \lambda + \sum_{n=1}^{\infty} a_n z^n \in \mathscr{S}_\lambda$,

$$\lambda + \sum_{n=1}^{\infty} |a_n| r^n \leqslant 1, \qquad r \in [0, \mathfrak{B}_0(\lambda)].$$

(b) For each $r \in (\mathfrak{B}_0(\lambda), 1)$, there is an $f = \lambda + \sum_{n=1}^{\infty} a_n z^n \in \mathscr{S}_\lambda$ such that

$$\lambda + \sum_{n=1}^{\infty} |a_n| r^n > 1.$$

As with Proposition 8.1.10,

$$\mathfrak{B}_0(\lambda) = \inf_{f \in \mathscr{S}_\lambda} c(f)$$

and $\mathfrak{B}_0(\lambda)$ is unique solution to $\mathfrak{M}(\mathscr{S}_\lambda, r) = 1$. Since $\mathfrak{B}_0 = \frac{1}{3}$ and $\mathscr{S}_\lambda \subseteq \mathscr{S}$, we conclude that

$$\mathfrak{B}_0(\lambda) \geqslant \frac{1}{3}, \qquad \lambda \in [0, 1).$$

Furthermore, as a consequence of the definitions,

$$\mathfrak{B}_0 = \inf_{0 \leqslant \lambda < 1} \mathfrak{B}_0(\lambda).$$

After developing more tools, we will estimate $\mathfrak{B}_0(\lambda)$ in Corollary 8.6.8 below.

In Corollary 8.2.6 we showed, for the classical Bohr radius $\mathfrak{B}_0 = \frac{1}{3}$, that there is no $f \in \mathscr{S}$ for which $\mathfrak{B}_0 = c(f)$. In other words, there is no $f = \sum_{n=0}^{\infty} a_n z^n \in \mathscr{S}$ for which

$$\sum_{n=0}^{\infty} |a_n| \mathfrak{B}_0^n = 1.$$

For the generalized Bohr radius $\mathfrak{B}_0(\lambda)$, the story is different.

Theorem 8.4.2 *For each* $\lambda \in [0, 1)$*, there is an* $f = \lambda + \sum_{n=1}^{\infty} a_n z^n \in \mathscr{S}_\lambda$ *such that* $\mathfrak{B}_0(\lambda) = c(f)$*; that is*

$$\lambda + \sum_{n=1}^{\infty} |a_n| \mathfrak{B}_0(\lambda)^n = 1.$$

Proof By (8.1.2) and (8.1.8),

$$\mathfrak{B}_0(\lambda) \leqslant \frac{1}{1 + 2\lambda} < 1.$$

Since

$$\mathfrak{B}_0(\lambda) = \inf_{f \in \mathscr{S}_\lambda} c(f),$$

there is a sequence $f_n \in \mathscr{S}_\lambda$ such that

$$\mathfrak{B}_0(\lambda) \leqslant c(f_n) \leqslant \min \left\{ \mathfrak{B}_0(\lambda) + \frac{1}{n}, \frac{1}{1 + 2\lambda} \right\}. \tag{8.4.3}$$

Because $f_n \in \mathscr{S}_\lambda \subseteq \mathscr{S}$, we see that $\{f_n : n \geqslant 1\}$ is a normal family. By Montel's theorem, it has a subsequence that converges uniformly on compact subsets of $\mathbb{D}$. Without loss of generality, we may assume that f_n itself converges to f uniformly on compact subsets of $\mathbb{D}$.

First, $f(0) = \lim_{n\to\infty} f_n(0) = \lambda$ and hence $f \in \mathscr{S}_\lambda$. Second, writing

$$f_n(z) = \lambda + \sum_{k=1}^{\infty} a_k(f_n) z^k \quad \text{and} \quad f(z) = \lambda + \sum_{k=1}^{\infty} a_k(f) z^k,$$

the Cauchy integral formula confirms that

$$\lim_{n\to\infty} a_k(f_n) = a_k(f), \qquad k \geqslant 1. \tag{8.4.4}$$

Third,

$$\lambda + \sum_{k=1}^{\infty} |a_k(f_n)| c(f_n)^k = 1,$$

and the terms of the series are dominated by

$$\sum_{k=1}^{\infty} \left(\frac{1}{1+2\lambda} \right)^k < \infty.$$

Note the use of the fact that $|a_n(f_n)| \leqslant 1$ since $f_n \in \mathscr{S}$: see Exercise 8.5. Therefore, by (8.4.3) and (8.4.4) and the discrete version of the dominated convergence theorem,

$$\lambda + \sum_{k=1}^{\infty} |a_k(f)| \mathfrak{B}_0(\lambda)^k = 1. \tag{8.4.5}$$

An important observation is that, since $\lambda < 1$, (8.4.5) implies that there is a k with $a_k(f) \neq 0$. Hence, f is not a constant function and thus $m(f, r)$ is strictly increasing. Therefore, (8.4.5) also implies that $c(f) = \mathfrak{B}_0(\lambda)$. □

If one attempts to adapt the procedure above for $\mathscr{S}$, then the proof works up to (8.4.5). But $a_k(f) = 0$ for all $k \geqslant 1$ and $|a_0(f)| = 1$. Hence we cannot proceed. In fact, we have Corollary 8.2.6 which says that f does not exist!

We call functions that satisfy the conclusion of Theorem 8.4.2 *extremal functions*. Later on, we will show that they are disk automorphisms.

The local Bohr coefficients $\mathfrak{B}_0(\lambda)$ can be generalized in the following way.

Definition 8.4.6 Fix any integer $k \geqslant 0$ and $\lambda \in [0, 1)$ and define $\mathfrak{B}_k(\lambda)$ to be the number in $[0, 1]$ that satisfies the following conditions.

(a) For all $f = \lambda z^k + \sum_{n=1}^{\infty} a_n z^{n+k} \in z^k \mathscr{S}_\lambda$,

$$\lambda r^k + \sum_{n=1}^{\infty} |a_n| r^{n+k} \leqslant 1, \qquad r \in [0, \mathfrak{B}_0(\lambda)].$$

(b) For each $r \in (\mathfrak{B}_0(\lambda), 1)$, there is an $f = \lambda z^k + \sum_{n=1}^{\infty} a_n z^{n+k} \in z^k \mathscr{S}_\lambda$ with

$$\lambda r^k + \sum_{n=1}^{\infty} |a_n| r^{n+k} > 1.$$

As in previous situations, we define

$$\mathfrak{B}_k(\lambda) := \inf_{f \in \mathscr{S}_\lambda} c(z^k f)$$

and $\mathfrak{B}_k(\lambda)$ is the unique solution of the equation $r^k \mathfrak{M}(\mathscr{S}_\lambda, r) = 1$. Furthermore,

$$\mathfrak{B}_k = \inf_{0 \leqslant \lambda < 1} \mathfrak{B}_k(\lambda), \qquad k \geqslant 0,$$

and

$$\mathfrak{B}_k(0) = \mathfrak{B}_{k+1}, \qquad k \geqslant 0.$$

We did not consider $\mathfrak{B}_k(1)$ for a good reason. Indeed, $\mathscr{S}_1$ contains only the constant function 1 and hence there is little to say.

8.5 Estimates of Landau and Bombieri

For $p \in (0, \infty)$, $r \in [0, 1)$, and an analytic function f on $\mathbb{D}$, let

$$f_r(e^{i\theta}) = f(re^{i\theta})$$

be a dilation of f and let

$$\|f_r\|_p = \left(\int_0^{2\pi} |f(re^{i\theta})|^p \frac{d\theta}{2\pi} \right)^{\frac{1}{p}}.$$

If $p \in [1, \infty)$, then $\lim_{r \to 1^-} \|f_r\|_p$ gives rise to the norm on the Hardy space H^p. This is entirely analogous to the introduction of the space H^2; see Appendix A.4. For $p \in (0, 1)$, the expression $\lim_{r \to 1^-} \|f_r\|_p$ no longer defines a norm because it fails to satisfy the triangle inequality.

Lemma 8.5.1 (Landau [92]) *Suppose that f is analytic on $|z| < R$ and has at least one zero there. If z_1 is a zero with minimal modulus, then for $r \in (|z_1|, R)$ and any $p \in (0, \infty)$,*

$$\frac{r|f(0)|}{\|f_r\|_p} \leqslant |z_1|.$$

Proof If $z_1 = 0$, then $f(0) = 0$ and the result is trivial. Consequently, we may assume that $f(0) \neq 0$ so that $z_1 \neq 0$. Let $z_1, z_2, \ldots, z_n$ be the zeros of f in $|z| < r$, ordered so that

$$0 < |z_1| \leqslant |z_2| \leqslant \cdots \leqslant |z_n| < r.$$

By Jensen's formula (Theorem A.5.1),

$$\log|f(0)| = \sum_{k=1}^{n} \log\left(\frac{|z_k|}{r}\right) + \int_0^{2\pi} \log|f(re^{i\theta})| \frac{d\theta}{2\pi}.$$

Each term in the sum is negative and hence

$$\log|f(0)| \leqslant \log\left(\frac{|z_1|}{r}\right) + \int_0^{2\pi} \log|f(re^{i\theta})| \frac{d\theta}{2\pi},$$

which we rewrite as

$$\frac{r|f(0)|}{|z_1|} \leqslant \exp\left(\int_0^{2\pi} \log|f(re^{i\theta})| \frac{d\theta}{2\pi}\right).$$

Take the pth power of both sides of the inequality above and then apply Jensen's inequality (Theorem A.5.2) to the right-hand side of

$$\left(\frac{r|f(0)|}{|z_1|}\right)^p \leqslant \exp\left(\int_0^{2\pi} \log|f(re^{i\theta})|^p \frac{d\theta}{2\pi}\right)$$

and deduce that

$$\left(\frac{r|f(0)|}{|z_1|}\right)^p \leqslant \int_0^{2\pi} |f(re^{i\theta})|^p \frac{d\theta}{2\pi}.$$

The result now follows upon rearranging the terms and taking pth roots. □

To effectively use Lemma 8.5.1 when $p = 2$, we need to estimate

$$\|f_r\|_2 = \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi}.$$

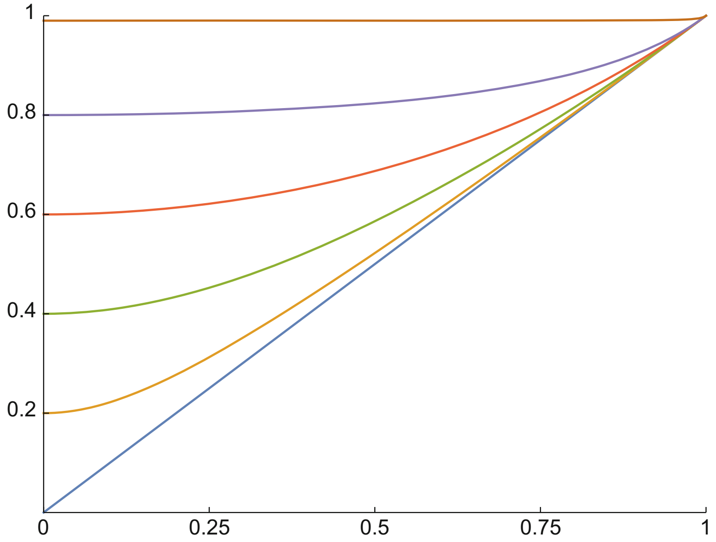


Fig. 8.2 The upper bound from the right-hand side of (8.5.3) (as a function of r) plotted for the values $\lambda = 0, 0.2, 0.4, 0.6, 0.8, 0.99$

This is done in the following lemma.

Lemma 8.5.2 (Bombieri [10]) *Let $f \in \mathscr{S}_\lambda$. Then for each $r \in [0, 1)$,*

$$\|f_r\|_2 \leqslant \left(\frac{r^2 + \lambda^2 - 2r^2\lambda^2}{1 - r^2\lambda^2} \right)^{\frac{1}{2}}. \tag{8.5.3}$$

Moreover, the following statements are equivalent (Fig. 8.2).

(a) The equality holds in (8.5.3) for some $r \neq 0$.
(b) The equality holds in (8.5.3) for all values of $r \in [0, 1)$.
(c) The function f is a disk automorphism

$$f(z) = \frac{\lambda - e^{i\alpha}z}{1 - \lambda e^{i\alpha}z},$$

in which α is an arbitrary real constant.

Proof Let b_λ denote the disk automorphism

$$b_\lambda(z) = \frac{\lambda - z}{1 - \lambda z}$$

and note that $b_\lambda \in \mathscr{S}_\lambda$. Then

$$g = b_\lambda \circ f$$

satisfies $g(0) = 0$, which means $g \in \mathscr{S}_0$. By the Schwarz Lemma (Lemma 1.1.1),

$$|g(z)| \leqslant |z|, \qquad z \in \mathbb{D}, \tag{8.5.4}$$

with equality for some $z \neq 0$ if and only if $g(z) = e^{i\alpha}z$ for some real α. Write (8.5.4) in terms of f as

$$|f(re^{i\theta}) - \lambda| \leqslant r|1 - \lambda f(re^{i\theta})|, \qquad r \in [0, 1). \tag{8.5.5}$$

Parseval's formula (A.4.3) provides

$$\frac{1}{2\pi}\int_0^{2\pi} |\lambda - f(re^{i\theta})|^2\, d\theta = |a_1|^2 r^2 + |a_2|^2 r^4 + \cdots = \|f_r\|_2^2 - \lambda^2. \tag{8.5.6}$$

On the other hand, using the same technique,

$$\frac{1}{2\pi}\int_0^{2\pi} |1 - \lambda f(re^{i\theta})|^2\, d\theta = 1 - 2\lambda^2 + \lambda^2\|f_r\|_2^2. \tag{8.5.7}$$

Then (8.5.5), (8.5.6), and (8.5.7) reveal that

$$\|f_r\|_2^2 - \lambda^2 \leqslant r^2(1 - 2\lambda^2 + \lambda^2\|f_r\|_2^2),$$

from which the desired result follows. Equality holds in the preceding for some $r > 0$ if and only if equality holds in (8.5.5) for almost all $re^{i\theta}$. This occurs if and only if $g(z) = e^{i\alpha}z$ for some real α. This leads to the proposed formula for f as a disk automorphism. □

8.6 A Theorem of Bombieri and Ricci

We are now ready to find the precise formula for $\mathfrak{M}(S_\lambda, r)$ for certain values of r and $\mathfrak{B}_0(\lambda)$ for $\frac{1}{2} \leqslant \lambda < 1$. We also provide estimates for the remaining values of r and λ; see Fig. 8.3.

Theorem 8.6.1 (Bombieri [10]) *If* $0 \leqslant \lambda < 1$*, then*

$$\mathfrak{M}(\mathscr{S}_\lambda, r) = \frac{\lambda + (1 - 2\lambda^2)r}{1 - \lambda r}, \qquad r \in [0, \lambda],$$

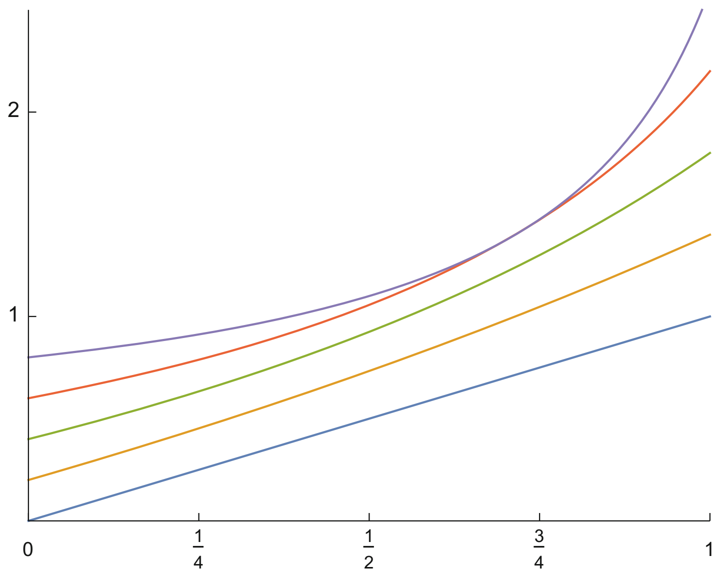


Fig. 8.3 Graphs of $\frac{\lambda+(1-2\lambda^2)r}{1-\lambda r}$ for the values $\lambda = 0, 0.2, 0.4, 0.6, 0.8$. For $r \in [0, \lambda]$, this quantity equals $\mathfrak{M}(\mathscr{S}_\lambda, r)$; for $r \in (\lambda, 1)$, it is a lower bound for $\mathfrak{M}(\mathscr{S}_\lambda, r)$

and

$$\frac{\lambda + (1 - 2\lambda^2)r}{1 - \lambda r} \leqslant \mathfrak{M}(\mathscr{S}_\lambda, r) \leqslant \lambda + r\left(\frac{1-\lambda^2}{1-r^2}\right)^{\frac{1}{2}}, \qquad r \in [\lambda, 1).$$

Proof To establish the lower bound, consider the disk automorphism

$$f(z) = \frac{\lambda - e^{i\alpha}z}{1 - \lambda e^{i\alpha}z}, \tag{8.6.2}$$

in which α is real. Then $f \in \mathscr{S}_\lambda$,

$$f(z) = \lambda + \sum_{n=1}^{\infty} (\lambda^2 - 1)\lambda^{n-1} e^{in\alpha} z^n, \qquad z \in \mathbb{D},$$

and

$$m(f, r) = \lambda + \sum_{n=1}^{\infty} (1 - \lambda^2)\lambda^{n-1} r^n = \frac{\lambda + (1 - 2\lambda^2)r}{1 - \lambda r}, \qquad r \in [0, 1).$$

Consequently,

$$\mathfrak{M}(\mathscr{S}_\lambda, r) \geqslant \frac{\lambda + (1 - 2\lambda^2)r}{1 - \lambda r}, \qquad r, \lambda \in [0, 1). \tag{8.6.3}$$

Finding functions f that yield equality in (8.6.3) is more delicate. Let

$$f = \lambda + \sum_{n=1}^{\infty} a_n z^n \in \mathscr{S}_\lambda$$

be nonconstant and define

$$g(z) = \lambda + \sum_{n=1}^{\infty} |a_n| z^n.$$

Analogous to the radius $c = c(f)$ defined in Sect. 8.1, let $c_\sigma = c_\sigma(f)$ be the point in $(0, 1)$ for which $g(c_\sigma) = \sigma$. Note that we need to assume that $\sigma \geqslant \lambda$. One can also see that $c = c_1$. Since the coefficients of $g(z)$ are all nonnegative, $|g(z)| < \sigma$ for all $|z| < c_\sigma$. Thus, we can apply Lemma 8.5.1 to the auxiliary function $h = \sigma - g$. Note that $h(c_\sigma) = 0$ and h has no zeros in $|z| < c_\sigma$. For $c_\sigma < r < 1$, Lemma 8.5.1 implies that

$$c_\sigma \geqslant \frac{(\sigma - \lambda)r}{\|h_r\|_2}. \tag{8.6.4}$$

The parameter $p = 2$ has the advantage that we can relate the L^2-norm of h_r with that of f_r. In fact,

$$\begin{aligned}
\|h_r\|_2^2 &= \frac{1}{2\pi} \int_0^{2\pi} |h(re^{i\theta})|^2 \, d\theta \\
&= (\sigma - \lambda)^2 + |a_1|^2 r^2 + \cdots \\
&= \sigma^2 - 2\lambda\sigma + (\lambda^2 + |a_1|^2 r^2 + \cdots) \\
&= \sigma^2 - 2\lambda\sigma + \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta \\
&= \sigma^2 - 2\lambda\sigma + \|f_r\|_2^2.
\end{aligned}$$

Therefore, by (8.6.4),

$$c_\sigma \geqslant \frac{(\sigma - \lambda)r}{\left(\sigma^2 - 2\lambda\sigma + \|f_r\|_2^2\right)^{\frac{1}{2}}}.$$

Up to now, we know that the inequality above is valid whenever $r \in (c_\sigma, 1)$. However, since $\|f_r\|_2 \geqslant |f(0)| = \lambda$, for any $r \in [0, 1)$ we have

$$\frac{(\sigma - \lambda)r}{\left(\sigma^2 - 2\lambda\sigma + \|f_r\|_2^2\right)^{\frac{1}{2}}} \leqslant r.$$

In particular, for $r \in [0, c_\sigma]$, we see that

$$c_\sigma \geqslant r \geqslant \frac{(\sigma - \lambda)r}{\left(\sigma^2 - 2\lambda\sigma + \|f_r\|_2^2\right)^{\frac{1}{2}}}.$$

Therefore, for each $f \in \mathscr{S}_\lambda$, the estimate

$$c_\sigma(f) \geqslant \frac{(\sigma - \lambda)r}{\left(\sigma^2 - 2\lambda\sigma + \|f_r\|_2^2\right)^{\frac{1}{2}}} \tag{8.6.5}$$

holds for all $r \in [0, 1)$. Applying the upper estimate in Lemma 8.5.2 gives us

$$c_\sigma(f) \geqslant (\sigma - \lambda)r \left[\sigma^2 - 2\lambda\sigma + \frac{r^2 + \lambda^2 - 2r^2\lambda^2}{1 - r^2\lambda^2}\right]^{-\frac{1}{2}}. \tag{8.6.6}$$

We have freedom in the choice of $r \in [0, 1)$ to obtain a good lower bound. By continuity, any value of $r \in [0, 1]$ is acceptable. At this point, we need to consider two cases.

Case I, $\sigma \in [\lambda, 2\lambda]$ The optimal result is obtained when

$$r = \left(\frac{\sigma - \lambda}{\lambda(1 - 2\lambda^2 + \lambda\sigma)}\right)^{\frac{1}{2}}.$$

The restriction $\sigma \in [\lambda, 2\lambda]$ ensures that $r \in [0, 1]$. By (8.6.6), this choice leads to the lower bound

$$c_\sigma(f) \geqslant \frac{\sigma - \lambda}{1 + \lambda\sigma - 2\lambda^2}.$$

Therefore, for each $f \in \mathscr{S}_\lambda$,

$$m\left(f, \frac{\sigma - \lambda}{1 + \lambda\sigma - 2\lambda^2}\right) \leqslant \sigma, \qquad \sigma \in [\lambda, 2\lambda].$$

If we put

$$s = \frac{\sigma - \lambda}{1 + \lambda\sigma - 2\lambda^2}$$

in the above we obtain

$$m(f, s) \leqslant \frac{\lambda + (1 - 2\lambda^2)s}{1 - \lambda s}, \qquad s \in [0, \lambda]. \tag{8.6.7}$$

Note that the restriction $\sigma \in [\lambda, 2\lambda]$ implies $s \in [0, \lambda]$ and thus, by (8.6.3), we obtain the precise formula for $\mathfrak{M}(\mathscr{S}_\lambda, s)$ for $s \in [0, \lambda]$.

Case II, $\sigma \in [2\lambda, \infty)$ The optimal radius is $r = 1$. Hence, by (8.6.6), we obtain the rough estimate

$$c_\sigma(f) \geqslant \frac{\sigma - \lambda}{(1 + \sigma^2 - 2\lambda\sigma)^{\frac{1}{2}}}.$$

Therefore, for each $f \in \mathscr{S}_\lambda$,

$$m\left(f, \frac{\sigma - \lambda}{(1 + \sigma^2 - 2\lambda\sigma)^{\frac{1}{2}}}\right) \leqslant \sigma, \qquad \sigma \geqslant 2\lambda.$$

Put

$$s = \frac{\sigma - \lambda}{(1 + \sigma^2 - 2\lambda\sigma)^{\frac{1}{2}}}$$

in the above to obtain the desired result. □

Corollary 8.6.8 (Bombieri [10], Ricci [117])

(a) $\mathfrak{B}_0(\lambda) = \dfrac{1}{1 + 2\lambda}$ *for* $\frac{1}{2} \leqslant \lambda < 1$.

(b) $\left(\dfrac{1 - \lambda}{2}\right)^{\frac{1}{2}} \leqslant \mathfrak{B}_0(\lambda) \leqslant \dfrac{1}{1 + 2\lambda}$ *for* $\dfrac{\sqrt{2} - 1}{2} \leqslant \lambda \leqslant \dfrac{1}{2}$.

(c) $\left(\dfrac{1 - \lambda}{2}\right)^{\frac{1}{2}} \leqslant \mathfrak{B}_0(\lambda) \leqslant \dfrac{1}{\sqrt{2}}$ *for* $0 \leqslant \lambda \leqslant \dfrac{\sqrt{2} - 1}{2}$.

Moreover, when $\lambda \in [\frac{1}{2}, 1)$, the extremal functions from Theorem 8.4.2 are precisely the disk automorphisms (8.6.2).

Proof *Case I, $\lambda \in [\frac{1}{2}, 1)$*: According to the equality in Theorem 8.6.1, the equation $\mathfrak{M}(\mathscr{S}_\lambda, r) = 1$ is equivalent to

$$\frac{\lambda + (1 - 2\lambda^2)r}{1 - \lambda r} = 1.$$

Solve for r and obtain

$$r = \mathfrak{B}_0(\lambda) = \frac{1}{1+2\lambda}. \tag{8.6.9}$$

In order for $r \in [0, \lambda]$ to occur, we must have

$$\frac{1}{1+2\lambda} \leqslant \lambda,$$

which leads to the imposed condition $\frac{1}{2} \leqslant \lambda < 1$.

We now discuss the extremal functions. In the first place, the previous paragraph shows that the disk automorphisms from (8.6.2) are extremal functions (even when $\lambda = \frac{1}{2}$). Now, suppose that f is an extremal function for this case; that is,

$$f \in \mathscr{S}_\lambda \qquad \text{and} \qquad c(f) = \frac{1}{1+2\lambda}, \qquad \lambda \in (\tfrac{1}{2}, 1).$$

Hence by (8.6.5),

$$\frac{1}{1+2\lambda} \geqslant \frac{(1-\lambda)r}{\left(1-2\lambda+\|f_r\|_2^2\right)^{\frac{1}{2}}}, \qquad r \in [0, 1).$$

Rearranging the terms, we can write the inequality above as

$$\|f_r\|_2 \geqslant \left((1-\lambda)^2(1+2\lambda)^2 r^2 - 1 + 2\lambda\right)^{\frac{1}{2}}, \qquad r \in [0, 1).$$

For the specific radius $r = (\lambda + 2\lambda^2)^{-\frac{1}{2}} \in (0, 1)$, this inequality becomes

$$\|f_r\|_2 \geqslant \left(\frac{r^2+\lambda^2-2r^2\lambda^2}{1-r^2\lambda^2}\right)^{\frac{1}{2}}.$$

Hence by Lemma 8.5.2, equality holds and f is the suggested disk automorphism.

Case II, $\lambda \in [0, \frac{1}{2})$ We now appeal to the inequalities in Theorem 8.6.1. We see that if $\mathfrak{M}(S_\lambda, r) = 1$, then

$$\frac{\lambda + (1-2\lambda^2)r}{1-\lambda r} \leqslant 1 \leqslant \lambda + r\left(\frac{1-\lambda^2}{1-r^2}\right)^{\frac{1}{2}}.$$

Solving for r gives

$$\left(\frac{1-\lambda}{2}\right)^{\frac{1}{2}} \leqslant r \leqslant \frac{1}{1+2\lambda}.$$

In this case $r \in [\lambda, 1)$. Hence, we need to assume that

$$\left(\frac{1-\lambda}{2}\right)^{\frac{1}{2}} \geqslant \lambda,$$

which leads to $0 \leqslant \lambda \leqslant \frac{1}{2}$. The lower bound is established, but we can improve the upper bound for $\lambda \leqslant \frac{\sqrt{2}-1}{2}$.

Case III, $0 \leqslant \lambda \leqslant \frac{\sqrt{2}-1}{2}$ Let

$$f(z) = \lambda + (1-\lambda)z\frac{a-z}{1-az}, \tag{8.6.10}$$

in which a is a free parameter in $(0, 1)$. Then $f \in \mathscr{S}_\lambda$ and, by (8.1.2),

$$f(z) = \lambda + (1-\lambda)az + (1-\lambda)(a^2-1)\sum_{n=2}^{\infty} a^{n-2}z^n. \tag{8.6.11}$$

Hence

$$\begin{aligned} m(f,r) &= \lambda + (1-\lambda)ar + (1-\lambda)(1-a^2)\sum_{n=2}^{\infty} a^{n-2}r^n \\ &= \lambda + (1-\lambda)\frac{ar + (1-2a^2)r^2}{1-ar}. \end{aligned}$$

Consequently, $c = c(f)$, which is obtained via the equation $m(f, c) = 1$, satisfies

$$(1-2a^2)c^2 - 2ac - 1 = 0. \tag{8.6.12}$$

The smallest c is obtained when $a = \frac{1}{\sqrt{2}}$ (see Exercise 8.6), which incidentally gives

$$c(f) = \frac{1}{\sqrt{2}}. \tag{8.6.13}$$

Therefore, again recalling that

$$\mathfrak{B}_0(\lambda) = \inf_{f \in \mathscr{S}_\lambda} c(f),$$

we conclude that

$$\mathfrak{B}_0(\lambda) \leqslant \frac{1}{\sqrt{2}}, \qquad \lambda \in [0, 1).$$

Compared with the bound $1/(1+2\lambda)$, this new bound is better if $\lambda \leqslant \frac{\sqrt{2}-1}{2}$. □

Corollary 8.6.8 implies $\mathfrak{B}_0(0) = \frac{1}{\sqrt{2}}$. However, $\mathfrak{B}_1 = \mathfrak{B}_0(0)$, and thus we obtain the second proof of that $\mathfrak{B}_1 = \frac{1}{\sqrt{2}}$.

From Lemma 8.1.13, we have the lower bound

$$\mathfrak{M}(\mathscr{S}, r) \geqslant \frac{3 - \sqrt{8(1-r^2)}}{r}, \qquad r \in [\tfrac{1}{3}, 1]. \tag{8.6.14}$$

The estimates in Theorem 8.6.1 enable us to show that the preceding is, in fact, an equality for $r \in [\frac{1}{3}, \frac{1}{\sqrt{2}}]$.

Theorem 8.6.15 (Bombieri [10])

$$\mathfrak{M}(\mathscr{S}, r) = \frac{3 - \sqrt{8(1-r^2)}}{r}, \quad r \in [\tfrac{1}{3}, \tfrac{1}{\sqrt{2}}].$$

Proof In light of (8.6.14), we need only establish that $\mathfrak{M}(\mathscr{S}, r)$ is at most the given quantity for r in the given range. Fix $r \in [\frac{1}{3}, \frac{1}{\sqrt{2}}]$. Since

$$\mathfrak{M}(\mathscr{S}, r) = \sup_{0 \leqslant \lambda < 1} \mathfrak{M}(\mathscr{S}_\lambda, r),$$

the estimates in Theorem 8.6.1 yield

$$\mathfrak{M}(\mathscr{S}, r) \leqslant \max\{A, B\},$$

in which

$$A = \sup_{r \leqslant \lambda < 1} \frac{\lambda + (1 - 2\lambda^2)r}{1 - \lambda r} \qquad \text{and} \qquad B = \sup_{0 \leqslant \lambda \leqslant r} \lambda + r\left(\frac{1-\lambda^2}{1-r^2}\right)^{\frac{1}{2}}.$$

One confirms (see Exercise 8.7) that

$$A = \frac{3 - \sqrt{8(1-r^2)}}{r}, \qquad B = \frac{1}{\sqrt{1-r^2}}, \tag{8.6.16}$$

and $A \geqslant B$ for $r \in [\frac{1}{3}, \frac{1}{\sqrt{2}}]$. □

A formula for $\mathfrak{M}(\mathscr{S}, r)$ for all values of $r \in [0, 1)$ is still unknown. Currently we know that

$$\mathfrak{M}(\mathscr{S}, r) = 1, \qquad r \in [0, \tfrac{1}{3}],$$

and

$$\mathfrak{M}(\mathscr{S}, r) = \frac{3 - \sqrt{8(1-r^2)}}{r}, \qquad r \in [\tfrac{1}{3}, \tfrac{1}{\sqrt{2}}].$$

For $r > \frac{1}{\sqrt{2}}$ a formula for $\mathfrak{M}(\mathscr{S}, r)$ is unknown. E. Bombieri and J. Bourgain [11] proved that

$$\frac{1}{\sqrt{1-r^2}} - c(\varepsilon)\left(\log \frac{1}{1-r}\right)^{\frac{3}{2+\varepsilon}} \leqslant \mathfrak{M}(\mathscr{S}, r) < \frac{1}{\sqrt{1-r^2}}, \qquad r \in [\tfrac{1}{\sqrt{2}}, 1). \tag{8.6.17}$$

In particular, this estimate implies that

$$\lim_{r \to 1^-} \mathfrak{M}(\mathscr{S}, r) = \infty,$$

which is an interesting and nontrivial fact.

8.7 Notes

Alternate Proofs

Alternate proofs and generalizations of the Bohr radius can be found in [110, 117, 120, 126, 134].

Other Bohr Inequalities

There are other Bohr inequalities for various subclasses of the Schur class. We refer the reader to the nice survey paper [104].

Harald Versus Nils

There is another "Bohr radius" from atomic physics due to Nils Bohr, the older brother of Harald Bohr.

Harald Bohr, Footballer

The mathematician Harald Bohr was a member of the Danish national football team. He won a silver medal at the 1908 Summer Olympics.

8.8 Exercises

8.1 Prove Proposition 8.1.10.

8.2 In the proof of Lemma 8.1.13, prove that the optimal value of

$$\frac{a + (1 - 2a^2)r}{1 - ar}$$

occurs when

$$a = \frac{2 - \sqrt{2(1 - r^2)}}{2r}.$$

8.3 Complete the details of the proof of Lemma 8.1.13.

8.4 In the proof of Wiener's inequality (Lemma 8.2.8), prove that the matrix representation of M_f is the Toeplitz matrix in (8.2.9).

8.5 Use Fourier coefficients (see (A.1.2)) along with Fatou's theorem (Theorem A.3.1) to verify that if

$$f = \sum_{n=0}^{\infty} a_n z^n$$

belongs to $\mathscr{S}$, then $|a_n| \leqslant 1$ for all n.

8.6 In (8.6.12) show that the smallest c in $(1 - 2a^2)c^2 - 2ac - 1 = 0$ is obtained when $a = \frac{1}{\sqrt{2}}$.

8.7 Confirm the identities in (8.6.16).

Chapter 9
Finite Blaschke Products and Group Theory

In this chapter we explore two connections between finite Blaschke products and finite group theory. For each finite Blaschke product B, we discuss the group of continuous maps $u : \mathbb{T} \to \mathbb{T}$ for which $B \circ u = B$ on $\mathbb{T}$. We also investigate conditions under which a finite Blaschke product B can be written as the composition of two non-automorphic finite Blaschke products. This is related to the monodromy group associated with B.

9.1 A Cyclic Subgroup

Let B be a finite Blaschke product of degree n. For each $w \in \mathbb{T}$, Theorem 3.4.10 says that the equation $B(z) = w$ has exactly n distinct solutions on $\mathbb{T}$. Thus, the sets $B^{-1}(\{w\})$ for $w \in \mathbb{T}$ form a partition of $\mathbb{T}$ and each set in the partition has exactly n elements. Write

$$B^{-1}(\{1\}) = \{e^{i\vartheta_1}, e^{i\vartheta_2}, \ldots, e^{i\vartheta_n}\},$$

in which the arguments are arranged so that

$$0 \leqslant \vartheta_1 < \vartheta_2 < \cdots < \vartheta_n < 2\pi.$$

Define $\vartheta_k \in [0, 2\pi)$ for $k \in \mathbb{Z}$ by

$$\vartheta_k = \vartheta_\ell \pmod{2\pi} \qquad \Longleftrightarrow \qquad k \equiv \ell \pmod{n},$$

where $k \equiv \ell \pmod{n}$ when $k - \ell \in n\mathbb{Z}$. For example, $\vartheta_{n+1} = \vartheta_1$ and $\vartheta_0 = \vartheta_n$.

As ζ moves once counterclockwise on $\mathbb{T}$, the Argument Principle shows that the image $B(\zeta)$ traverses the unit circle n times. As ζ passes from $e^{i\vartheta_k}$ to $e^{i\vartheta_{k+1}}$,

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_9

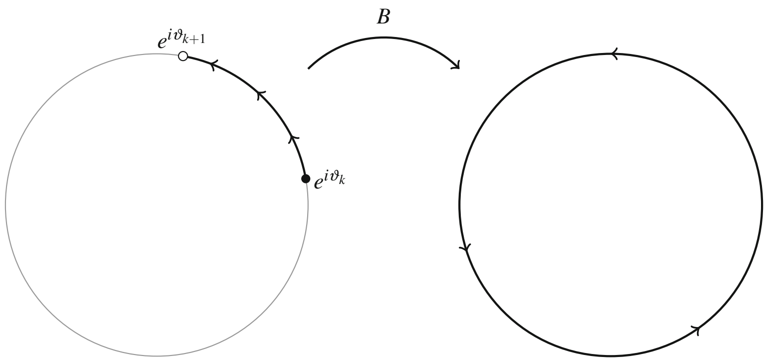


Fig. 9.1 B maps the arc subtended by $e^{i\vartheta_k}$ and $e^{i\vartheta_{k+1}}$ once around the unit circle

the image $B(\zeta)$ makes exactly one complete traversal of $\mathbb{T}$; see Fig. 9.1. Thus, B bijectively maps each of the arcs

$$[e^{i\vartheta_1}, e^{i\vartheta_2}), \quad [e^{i\vartheta_2}, e^{i\vartheta_3}), \ldots, [e^{i\vartheta_n}, e^{i\vartheta_{n+1}}),$$

onto $\mathbb{T}$. For each $k \in \mathbb{Z}$, define the bijective continuous function

$$\Phi_k : [e^{i\vartheta_k}, e^{i\vartheta_{k+1}}) \to \mathbb{T}, \qquad \Phi_k(e^{i\theta}) = B(e^{i\theta}). \tag{9.1.1}$$

This produces only n distinct functions since

$$\Phi_k \equiv \Phi_\ell \quad \Longleftrightarrow \quad k \equiv \ell \pmod{n}.$$

For two functions f and g on a set E, we use the notation $f \equiv g$ when $f(x) = g(x)$ for all $x \in E$. According to the definition of the arguments ϑ_k, we see that

$$\lim_{\substack{\theta \to \vartheta_k \\ \theta > \vartheta_k}} \Phi_k(e^{i\theta}) = \lim_{\substack{\theta \to \vartheta_{k+1} \\ \theta < \vartheta_{k+1}}} \Phi_k(e^{i\theta}) = 1 \tag{9.1.2}$$

and

$$B\big(\Phi_k^{-1}(e^{i\theta})\big) = e^{i\theta}, \quad e^{i\theta} \in \mathbb{T}. \tag{9.1.3}$$

Define an equivalence relation $\sim$ on $\mathbb{T}$ by

$$e^{i\theta_1} \sim e^{i\theta_2} \quad \Longleftrightarrow \quad B(e^{i\theta_1}) = B(e^{i\theta_2}).$$

Then

$$\{B^{-1}(\{w\}) : w \in \mathbb{T}\} \tag{9.1.4}$$

is the family of equivalence classes of $\sim$. Each of the n arcs described above contains exactly one element from each equivalence class. The following result shows that elements of a conjugacy class (9.1.4) are uniformly separated from each other.

Lemma 9.1.5 *If B is a finite Blaschke product, then there is $\delta > 0$ so that*

$$0 < |e^{i\theta} - e^{i\vartheta}| < \delta \quad \Longrightarrow \quad B(e^{i\theta}) \neq B(e^{i\vartheta}).$$

Proof Lemma 3.4.3, along with continuity and compactness, show that $|B'|$ is bounded away from zero on $\mathbb{T}$. Thus, by the Mean Value Theorem, there is a constant $C > 0$ so that $|B(e^{is}) - B(e^{it})| \geqslant C|s - t|$ for all s, t. □

Let $\mathscr{C}$ be the set of all continuous functions $u : \mathbb{T} \to \mathbb{T}$. This set, when endowed with the binary operation of function composition, is a semigroup. Indeed,

(a) $u_1, u_2 \in \mathscr{C} \implies u_1 \circ u_2 \in \mathscr{C}$;
(b) $(u_1 \circ u_2) \circ u_3 = u_1 \circ (u_2 \circ u_3)$ for all $u_1, u_2, u_3 \in \mathscr{C}$;
(c) $\mathrm{id} \in \mathscr{C}$.

Here id denotes the identity map on $\mathbb{T}$, which satisfies $u \circ \mathrm{id} = \mathrm{id} \circ u = u$ for each $u \in \mathscr{C}$. It is important to note that an arbitrary element of $\mathscr{C}$ need not be invertible under composition. For example, consider $u(z) = z^2$ (a branch of $\sqrt{z}$ cannot be defined on all of $\mathbb{T}$).

If B is a finite Blaschke product, then we may regard it as a (generally non-invertible) element of $\mathscr{C}$ and define

$$G_B := \{u \in \mathscr{C} : B \circ u = B\}.$$

A short argument shows that G_B is a sub-semigroup of $\mathscr{C}$. In fact, much more is true.

Theorem 9.1.6 (Cassier–Chalendar [18]) *Let B be a finite Blaschke product of degree n. Then G_B is a cyclic group of order n.*

Proof Consider the bijective mappings Φ_k defined in (9.1.1). For $k \in \mathbb{Z}$, define functions $u_k : \mathbb{T} \to \mathbb{T}$ by

$$u_k : [e^{i\vartheta_j}, e^{i\vartheta_{j+1}}) \to [e^{i\vartheta_{j+k}}, e^{i\vartheta_{j+k+1}}), \qquad u_k(e^{i\theta}) = \Phi_{j+k}^{-1}\big(\Phi_j(e^{i\theta})\big),$$

for $j \in \mathbb{Z}$ (see Fig. 9.2). Upon gluing these pieces together, we obtain a continuous bijection $u_k : \mathbb{T} \to \mathbb{T}$. Moreover, (9.1.2) implies that

$$u_k(e^{i\vartheta_j}) = e^{i\vartheta_{j+k}} \quad \text{and} \quad \lim_{\substack{\theta \to \vartheta_{j+1} \\ \theta < \vartheta_{j+1}}} u_k(e^{i\theta}) = e^{i\vartheta_{j+k+1}}.$$

Now observe that (9.1.3) ensures that

$$B\big(u_k(e^{i\theta})\big) = B\big(\Phi_{j+k}^{-1}(\Phi_j(e^{i\theta}))\big) = \Phi_j(e^{i\theta}) = B(e^{i\theta})$$

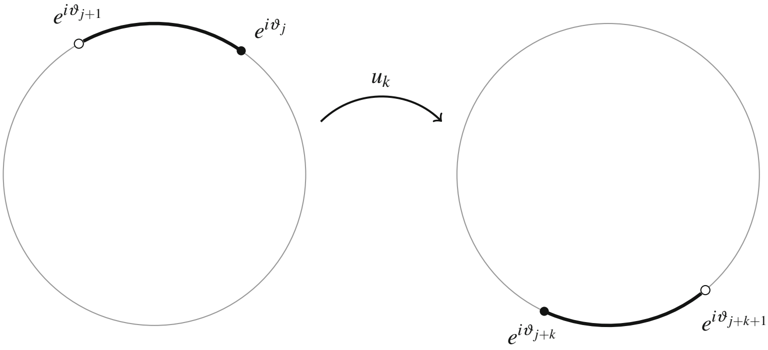


Fig. 9.2 The map u_k takes the half open arcs subtended by $e^{i\vartheta_j}, e^{i\vartheta_{j+1}}$ bijectively to the half open arc subtended by $e^{i\vartheta_{j+k}}, e^{i\vartheta_{j+k+1}}$

for each $e^{i\theta} \in \mathbb{T}$. In other words, by the construction above, we obtain n elements of G_B.

To further clarify the situation, let us make the following observations.

(a) $u_0 \equiv \text{id}$.
(b) $u_k \equiv u_\ell \iff k \equiv \ell \pmod{n}$.
(c) $u_k \equiv u_1 \circ u_1 \circ \cdots \circ u_1$ (k times).
(d) $u_k \circ u_\ell \equiv u_{k+\ell}$.
(e) naïvely speaking, we say that u_k shifts forward each of the arcs

$$[e^{i\vartheta_1}, e^{i\vartheta_2}), \quad [e^{i\vartheta_2}, e^{i\vartheta_3}), \ldots, \quad [e^{i\vartheta_n}, e^{i\vartheta_{n+1}}),$$

by k steps in such a way that it preserves the equivalence classes of $\sim$. The identity $u_k(\zeta) = u_k(\xi)$ implies that ζ and ξ belong to the same equivalence class of $\sim$.

These observations reveal that $\{u_0, u_1, \ldots, u_{n-1}\}$ is a cyclic subgroup of order n in G_B. We claim that this exhausts G_B. This fact is based on the following property: if $u, v \in G_B$ are such that $u(e^{i\theta_0}) = v(e^{i\theta_0})$ for some $e^{i\theta_0} \in \mathbb{T}$, then $u = v$. To verify this, let

$$E = \{e^{i\theta} \in \mathbb{T} : u(e^{i\theta}) = v(e^{i\theta})\}.$$

By assumption $e^{i\theta_0} \in E$. Since u and v are continuous functions, E is a closed subset of $\mathbb{T}$. By uniform continuity, there is a $\delta' > 0$ such that

$$\text{dist}(e^{i\theta}, E) < \delta' \quad \implies \quad |u(e^{i\theta}) - v(e^{i\theta})| < \delta,$$

in which $\delta > 0$ is the parameter introduced in Lemma 9.1.5. According to the definition of G_B, we have

$$B(u(e^{i\theta})) = B(v(e^{i\theta})) = B(e^{i\theta}), \qquad e^{i\theta} \in \mathbb{T}.$$

By Lemma 9.1.5, $u(e^{i\theta}) = v(e^{i\theta})$ at least for all $e^{i\theta}$ such that $\operatorname{dist}(e^{i\theta}, E) < \delta'$. This shows that E is also an open set, so $E = \mathbb{T}$.

To finish the proof, let $u \in G_B$. Then

$$B(u(e^{i\vartheta_1})) = B(e^{i\vartheta_1}) = 1,$$

and hence

$$u(e^{i\vartheta_1}) \in B^{-1}(\{1\}) = \{e^{i\vartheta_1}, e^{i\vartheta_2}, \dots, e^{i\vartheta_n}\}.$$

Suppose that $u(e^{i\vartheta_1}) = e^{i\vartheta_k}$ for some $1 \leqslant k \leqslant n$. If we rewrite this identity as $u(e^{i\vartheta_1}) = u_k(e^{i\vartheta_1})$, then the preceding observation shows that $u = u_k$. □

The following fact was stated and verified in the proof of Theorem 9.1.6.

Corollary 9.1.7 *Let B be a finite Blaschke product. Let $u : \mathbb{T} \to \mathbb{T}$ be a continuous function such that $B \circ u = B$. Suppose that there is an $e^{i\theta_0} \in \mathbb{T}$ so that $u(e^{i\theta_0}) = e^{i\theta_0}$. Then $u = \mathrm{id}$.*

9.2 Decomposable Finite Blaschke Products

We have seen in Theorem 3.6.2 that finite Blaschke products are closed under composition. Indeed, if C and D are finite Blaschke products, then $B = C \circ D$ is a finite Blaschke product with $\deg B = (\deg C)(\deg D)$. In this section we consider the following question.

Question 9.2.1 *When can a finite Blaschke product B be written as*

$$B = C \circ D,$$

in which C and D are finite Blaschke products of degree greater than one?

The restriction that both C and D are of degree greater than one avoids "trivial" decompositions such as

$$B = \phi \circ (\phi^{-1} \circ B) \quad \text{or} \quad B = (B \circ \phi) \circ \phi^{-1},$$

where $\phi \in \operatorname{Aut}(\mathbb{D})$ (which is a finite Blaschke product of degree one).

Definition 9.2.2 If $B = C \circ D$ in a nontrivial way, then B is *decomposable*. Otherwise, B is *indecomposable*.

Observe that if B is of prime degree, then B is indecomposable (Theorem 3.6.2).

The decomposability criterion covered here is a deep theorem of Ritt [119] (see Theorem 9.6.1 below) that classifies decomposability in terms of the monodromy group associated with B. Our treatment is based on unpublished notes of Carl Cowen [28] which he kindly agreed to let us use. Two other good sources for this are [108, 109].

Definition 9.2.3 If B is a finite Blaschke product of degree n, then B is in *normalized form* if

$$B(z) = z \prod_{k=2}^{n} \frac{\overline{a_k}}{|a_k|} \frac{a_k - z}{1 - \overline{a_k} z}, \tag{9.2.4}$$

in which $a_2, \ldots, a_n \in \mathbb{D}\backslash\{0\}$ are distinct. This is equivalent to the properties

$$B(0) = 0, \qquad B'(0) > 0, \qquad B(a) = 0 \implies B'(a) \neq 0. \tag{9.2.5}$$

Indeed, the first property follows from the factor of z in (9.2.4) and the last property follows from the simplicity of the zeros. The second property follows from the identity

$$B'(z) = z \frac{d}{dz}\Big(\prod_{k=2}^{n} \frac{\overline{a_k}}{|a_k|} \frac{a_k - z}{1 - \overline{a_k} z}\Big) + \prod_{k=2}^{n} \frac{\overline{a_k}}{|a_k|} \frac{a_k - z}{1 - \overline{a_k} z},$$

which implies that

$$B'(0) = \prod_{k=2}^{n} |a_k| > 0.$$

Since we will state our decomposability condition for a finite Blaschke product in normalized form, we first need to reduce the original decomposability problem, stated for general finite Blaschke products, to one for Blaschke products in normalized form.

Proposition 9.2.6 *If B is a finite Blaschke product, then there are $\alpha, \beta \in \mathbb{D}$ and $\xi \in \mathbb{T}$ such that $\widetilde{B} = \xi(\tau_\beta \circ B \circ \tau_\alpha)$ is a finite Blaschke product in normalized form. Moreover, B is decomposable if and only if $\widetilde{B}$ is decomposable.*

Proof If B is a finite Blaschke product of degree n, let

$$V = \{z \in \mathbb{D} : B'(z) = 0\}$$

denote the set of critical points of B and observe that V has cardinality $n - 1$ (Theorem 6.1.4). The set $B(V)$ is also finite and hence there is a $\beta \in \mathbb{D}\backslash B(V)$. Thus

$$B^{-1}(\{\beta\}) = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$$

(the fact that there are n pre-images follows from Theorem 3.4.10) with the property that

$$B(\alpha_j) = \beta \quad \text{and} \quad B'(\alpha_j) \neq 0. \tag{9.2.7}$$

Let $\alpha = \alpha_1$. For a unimodular constant ξ to be determined shortly, define

$$\widetilde{B}(z) = \xi(\tau_\beta \circ B \circ \tau_\alpha) \tag{9.2.8}$$

and notice that $\widetilde{B}$ is a finite Blaschke product of degree n that is a composition (pre- and post-) of B with disk automorphisms. This will be important in a moment.

Since $\tau_\alpha(0) = \alpha$, we have

$$\widetilde{B}(0) = \xi \frac{\beta - B(\alpha)}{1 - \overline{\beta} B(\alpha)} = 0$$

and hence $\widetilde{B}$ satisfies the first property in (9.2.5) of a normalized form.

A calculation with the quotient and chain rules yields

$$\widetilde{B}' = \xi \frac{B'(\tau_\alpha)\tau_\alpha'(-1 + |\beta|^2)}{(1 - \overline{\beta} B(\tau_\alpha))^2}. \tag{9.2.9}$$

Using the identity

$$\tau_\alpha'(z) = \frac{-1 + |\alpha|^2}{(1 - \overline{\alpha} z)^2},$$

we can substitute $z = 0$ into the expression above for $\widetilde{B}'$ to get

$$\widetilde{B}'(0) = \xi B'(\alpha) \frac{1 - |\alpha|^2}{1 - |\beta|^2} \neq 0$$

since $B'(\alpha) = B'(\alpha_1) \neq 0$ by (9.2.7). Now adjust the unimodular constant ξ so that $\widetilde{B}'(0) > 0$, which yields the second property in (9.2.5) of a normalized form.

Next, observe in the definition of $\widetilde{B}$ and (9.2.7) that the zeros of $\widetilde{B}$ are

$$\tau_\alpha^{-1}(\{\alpha_1, \alpha_2, \ldots, \alpha_n\}) = \{w_1, w_2, \ldots, w_n\}.$$

Note that $w_1 = 0$. From (9.2.9) we see that

$$\widetilde{B}'(w_j) = \xi B'(\tau_\alpha(w_j)) \frac{1 - |\alpha|^2}{(1 - \overline{\alpha} w_j)^2 (1 - |\beta|^2)}$$

$$= \xi B'(\alpha_j) \frac{1 - |\alpha|^2}{(1 - \overline{\alpha} w_j)^2 (1 - |\beta|^2)}$$

$$\neq 0$$

since $B'(\alpha_j) \neq 0$ from (9.2.7). Thus, $\widetilde{B}$ satisfies the third property in (9.2.5) of a normalized form.

To finish, we need to argue that B is decomposable if and only if $\widetilde{B}$ is decomposable. To this end, suppose that B_1, B_2 are finite Blaschke products with

$$B_1 = \phi \circ B_2 \circ \psi,$$

in which $\phi, \psi \in \operatorname{Aut}(\mathbb{D})$ and $B_2 = C \circ D$ is decomposable. Then, via the associativity of function composition,

$$B_1 = (\phi \circ C) \circ (D \circ \psi).$$

Moreover,

$$\deg(\phi \circ C) = \deg C \quad \text{and} \quad \deg(D \circ \psi) = \deg D.$$

A similar argument applies to B_2 and hence B_1 is decomposable if and only if B_2 is decomposable. Apply this fact to $B_1 = B$ and $B_2 = \widetilde{B}$ to complete the proof. □

Thus, in terms of whether or not a finite Blaschke product is decomposable, we can assume that it is in normalized form.

9.3 The Monodromy Group

For a finite Blaschke product B of degree n in normalized form (9.2.4) let

$$\mathscr{S}_B = \{w \in \mathbb{D} : w = B(z),\ B'(z) = 0\} \tag{9.3.1}$$

denote the set of *critical values* of B. Notice that $\mathscr{S}_B$ is the *image* of the set of critical points $\{z \in \mathbb{D} : B'(z) = 0\}$ of B. Since there are $n-1$ critical points in $\mathbb{D}$ (Theorem 6.1.4), $\mathscr{S}_B$ has at most $n-1$ points in $\mathbb{D}$. Now define

$$\widetilde{\mathscr{S}}_B := B^{-1}(\mathscr{S}_B) \tag{9.3.2}$$

and observe that $\widetilde{\mathscr{S}}_B$ contains at most $n(n-1)$ points of $\mathbb{D}$.

Consider the n-valued analytic function B^{-1} on $\mathbb{D}\backslash\mathscr{S}_B$. Since $0 \in \mathbb{D}\backslash\mathscr{S}_B$ by our normalizing assumption, B^{-1} has n branches

$$g_1, g_2, \ldots, g_n$$

at 0 where, for normalizing purposes, we number these branches so that

$$g_1(0) = 0.$$

Example 9.3.3 Let

$$B(z) = z\frac{\frac{1}{2} - z}{1 - \frac{z}{2}}$$

and observe that B is in normalized form. The only critical point (in $\mathbb{D}$) is

$$z = 2 - \sqrt{3} \approx 0.267949$$

and the critical value is

$$B(2 - \sqrt{3}) = 7 - 4\sqrt{3} \approx 0.0717968.$$

The two branches are

$$g_1(z) = \frac{1}{4}\left(z + 1 - \sqrt{z^2 - 14z + 1}\right) \text{ and } g_2(z) = \frac{1}{4}\left(z + 1 + \sqrt{z^2 - 14z + 1}\right).$$

Observe that $g_1(0) = 0$ while $g_2(0) = \frac{1}{2}$ (and that $0, \frac{1}{2}$ are the two zeros of B).

Let

$$L_B = \{\gamma : [0, 1] \to \mathbb{D}\backslash\mathscr{S}_B \ : \ \gamma \text{ is continuous}, \gamma(0) = \gamma(1) = 0\}$$

be the set of continuous closed curves in $\mathbb{D}\backslash\mathscr{S}_B$ that begin and end at the origin; see Fig. 9.3.

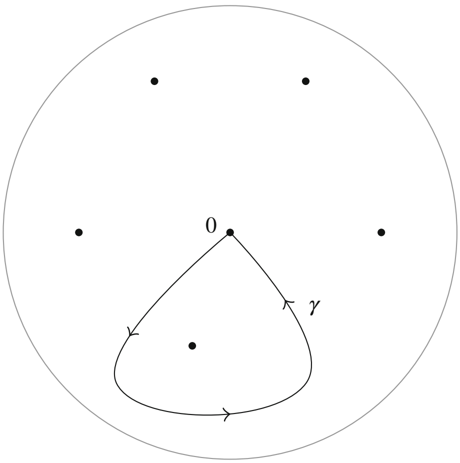


Fig. 9.3 A typical curve $\gamma \in L_B$. Notice how γ does not intersect the critical values of B

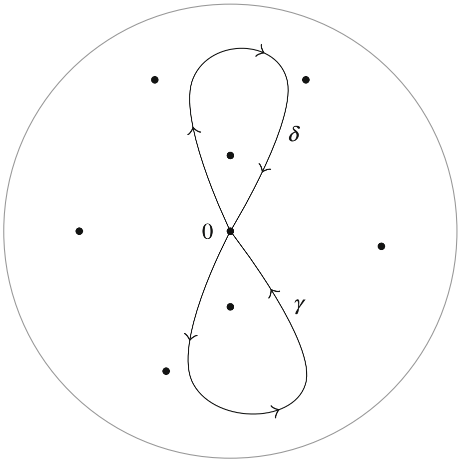


Fig. 9.4 The curve $\gamma \cdot \delta$ is obtained by following δ with γ

For $\gamma, \delta \in L_B$, let

$$\gamma \cdot \delta \tag{9.3.4}$$

be the element of L_B obtained by starting with δ and then continuing with γ; see Fig. 9.4. This is the usual product of curves used in algebraic topology.

If $\gamma \in L_B$, then g_1 can be analytically continued along γ and we let $\gamma^* g_1$ denote the final element of this continuation (see [100, Vol. III, Ch. 8] for a treatment of analytic continuation along arcs). Since $\gamma^* g_1$ must be one of the branches of B^{-1} at 0, we have

$$\gamma^* g_1 \in \{g_1, g_2, \ldots, g_n\}.$$

We can do an analogous construction to define the final elements

$$\gamma^* g_2, \gamma^* g_3, \ldots, \gamma^* g_n.$$

In each case,

$$\gamma^* g_j \in \{g_1, g_2, \ldots, g_n\}.$$

The alert reader might think that the definition of $\gamma^* g_j$, the final element of the continuation of g_j along γ, depends on the curve γ. It does not. Indeed, if $\gamma_1, \gamma_2 \in L_B$ with γ_1 homotopic in $\mathbb{D} \backslash \mathscr{S}_B$ to γ_2, then by the homotopy lemma [105], $\gamma_1^* g_j = \gamma_2^* g_j$. In other words, the definition of $\gamma^* g_j$ depends only on an element of the equivalence class of curves in L_B that are homotopic to γ in $\mathbb{D} \backslash \mathscr{S}_B$. We now think of γ^* as a function from the set of branches $\{g_1, g_2, \ldots, g_n\}$ to itself.

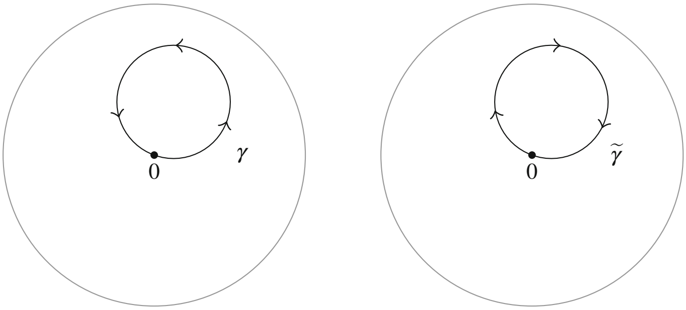


Fig. 9.5 The curve $\widetilde{\gamma}$ is the same as γ but traversed in the opposite direction

Definition 9.3.5 For a finite Blaschke product B in normal form, let

$$\mathscr{G}_B := \{\gamma^* : \gamma \in L_B\}.$$

For $\gamma \in L_B$, define $\widetilde{\gamma} \in L_B$ by the parameterization

$$\widetilde{\gamma}(t) = \gamma(1-t), \quad t \in [0, 1].$$

Observe that $\widetilde{\gamma}$ is just γ with the direction reversed; see Fig. 9.5. We use $\mathbf{e}$ to denote the element of L_B parameterized by

$$\mathbf{e}(t) = 0, \quad t \in [0, 1].$$

With the operation $\gamma \cdot \delta$ defined in (9.3.4), note that

$$\widetilde{\gamma} \cdot \gamma = \gamma \cdot \widetilde{\gamma} = \mathbf{e}.$$

We also see that

$$\mathbf{e}^* g_j = g_j, \qquad 1 \leqslant j \leqslant n,$$

and

$$(\gamma \cdot \delta)^* g_j = \gamma^*(\delta^* g_j), \qquad 1 \leqslant j \leqslant n.$$

Thus, $\mathscr{G}_B$ is a set with a well-defined binary operation $\gamma^* \circ \delta^*$ (which is also compatible with the operation $\gamma \cdot \delta$) and with an identity element $\mathbf{e}^*$. Moreover, the preceding also shows that

$$\widetilde{\gamma}^* \circ \gamma^* = \gamma^* \circ \widetilde{\gamma}^* = \mathbf{e}^*$$

and hence each element of $\mathscr{G}_B$ has an inverse. Consequently, $\mathscr{G}_B$ is a group, the *monodromy group* of B, and each element of $\mathscr{G}_B$ is a permutation of the branches. By equating the branches $\{g_1, g_2, \dots, g_n\}$ with the set $\{1, 2, \dots, n\}$ in the natural way, we see that $\mathscr{G}_B$ is isomorphic to a subgroup of the symmetric group S_n, the group of permutations of the set $\{1, 2, \dots, n\}$.

9.4 Examples of Monodromy Groups

Through the next several examples, we use a technique of Cowen [28] to "see" the monodromy group $\mathscr{G}_B$.

Example 9.4.1 The finite Blaschke product

$$b(z) = z^4$$

is not in normalized form since, among other things, it has a zero of degree 4 at the origin. We follow the recipe from the proof of Proposition 9.2.6. Indeed, we compute the critical points

$$V = \{z : b'(z) = 0\} = \{0\}$$

and choose $\frac{1}{2} \in \mathbb{D} \backslash b(0) = \mathbb{D} \backslash \{0\}$. To put b in normalized form, observe that since $b(\frac{1}{2}) = \frac{1}{16}$, we can use the formula in (9.2.8) to define

$$B(z) = \tau_{1/16} \circ b \circ \tau_{1/2} = \frac{\frac{1}{16} - b(\frac{1/2-z}{1-z/2})}{1 - \frac{1}{16} b(\frac{1/2-z}{1-z/2})}.$$

A computation shows that

$$B'(0) = \frac{32}{85} > 0$$

and that B has zeros at

$$z_1 = 0, \qquad z_2 = \frac{10}{17} - \frac{6i}{17}, \qquad z_3 = \frac{10}{17} + \frac{6i}{17}, \qquad z_4 = \frac{4}{5}.$$

Thus, B is in normalized form. One computes the set of critical *points*

$$\{z \in \mathbb{D} : B'(z) = 0\} = \{\tfrac{1}{2}\}$$

and the set of critical *values*

$$\mathscr{S}_B = \{w \in \mathbb{D} : w = B(z),\ B'(z) = 0\} = \{\tfrac{1}{16}\}.$$

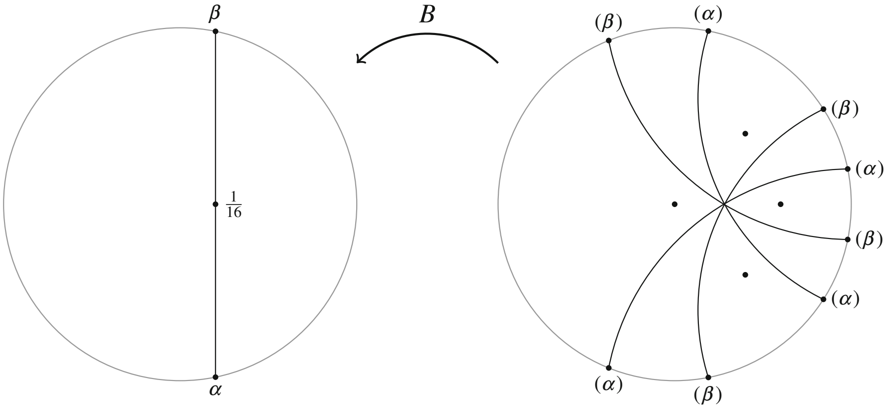


Fig. 9.6 (*Left*) The line δ that connects α and β and passes through the critical value $z = \frac{1}{16}$. (*Right*) The inverse image of $B^{-1}(\delta)$. The point on the right where all of the curves meet is the critical point $z = \frac{1}{2}$ which gets mapped to the critical value $\frac{1}{16}$ by B. If one travels on one of the triangular shaped regions that does not contain one of the zeros of B (marked by dots in the right hand image), from (α) to the center point $\frac{1}{2}$ to (β) and then along $\mathbb{T}$ back to α, then Rouché's theorem ensures that B maps this path from α through the critical value $\frac{1}{16}$ to β along $\mathbb{T}$ (to the right) back to α

Let us now compute $\mathscr{G}_B$ using [28]. The first step is to pick two different points α and β on $\mathbb{T}$ and a curve $\delta \subseteq \mathbb{D}^-$ that passes through α, β and the critical value $\frac{1}{16}$. For this particular case, we pick

$$\alpha = \frac{1}{16} - i\sqrt{1 - \frac{1}{16^2}}, \qquad \beta = \frac{1}{16} + i\sqrt{1 - \frac{1}{16^2}},$$

and let δ be the chord that connects α and β (which passes through $\frac{1}{16}$ but not 0); see Fig. 9.6.

Observe that $B^{-1}(\delta)$ is, at least locally, a curve except at the critical points of B, where $B^{-1}(\delta)$ will consist of intersecting curves; see Fig. 9.6 in which $(\alpha) = B^{-1}(\{\alpha\})$ and $(\beta) = B^{-1}(\{\beta\})$). The curve δ, a straight line in this case, divides $\mathbb{D}$ into two regions, one of which contains zero, while $B^{-1}(\delta)$ divides $\mathbb{D}$ into eight regions, four of which contain a zero of B; see Fig. 9.6.

There are several types of homotopy classes of curves $\gamma \in L_B$ one can consider when exploring γ^*. The first are the curves γ that do not loop around the critical value $z = \frac{1}{16}$. A representative example of such a curve is shown in Fig. 9.7. Notice how γ_1 starts at the origin, crosses δ between $z = \frac{1}{16}$ and $z = \beta$, loops around, crosses δ again between $z = \frac{1}{16}$ and β, before it returns to the origin. This means that $B^{-1}(\gamma_1)$ will start at the z_j, cross $B^{-1}(\delta)$ between the critical point $z = \frac{1}{2}$ and (β), turn around, recross $B^{-1}(\delta)$ between (β) and the critical point $z = \frac{1}{2}$, before it returns to z_j. With g_1, g_2, g_3, g_4 being the branches of B^{-1} with the understanding

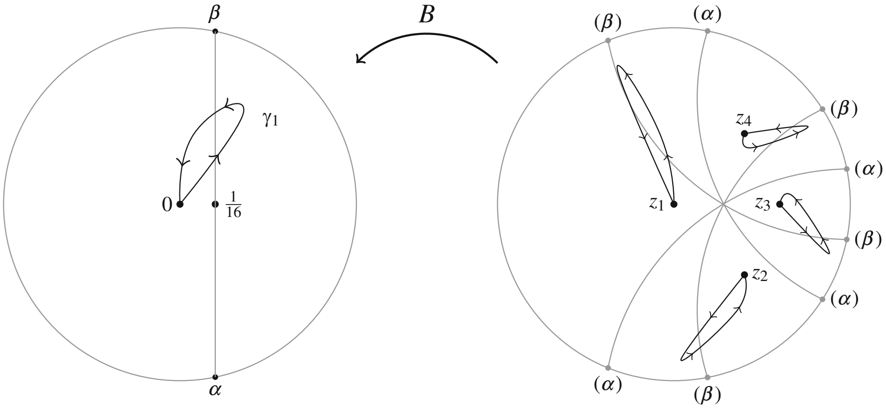


Fig. 9.7 (*Left*) A curve γ_1 that starts at the origin, crosses δ between $z = \frac{1}{16}$ and $z = \beta$, loops around, crosses δ again between $z = \frac{1}{16}$ and β, before returning to the origin. (*Right*) The inverse image curves for γ_1. Observe how γ_1 starts at 0, crosses δ (the straight line connecting α and β) between the critical value $z = \frac{1}{16}$ and β, turns and crosses δ again before it returns to 0. Notice how the image curves $B^{-1}(\gamma_1)$ start at the zeros of B, crosses $B^{-1}(\delta)$ between the critical point $z = \frac{1}{2}$ and (β), turn and cross $B^{-1}(\delta)$ again before they return to the respective zero of B

that $g_1(0) = 0$, we see that the final element $\gamma_1^* g_j$ is indeed g_j. Thus, we have

$$\gamma_1^* : \{g_1, g_2, g_3, g_4\} \to \{g_1, g_2, g_3, g_4\}, \qquad \gamma_1^*(g_j) = g_j.$$

In other words, γ_1^* is the identity permutation.

Now consider $\gamma_2 \in L_B$ which loops once around the critical value $z = \frac{1}{16}$; see Fig. 9.8. Notice how γ_2 first crosses δ between the critical value $\frac{1}{16}$ and α, turns, recrosses δ between β and the critical value, before it returns to the origin. With this in mind, observe how $B^{-1}(\gamma_2)$ starts off at z_j, crosses $B^{-1}(\delta)$ between (α) and the critical point $z = \frac{1}{2}$, crosses $B^{-1}(\delta)$ between (β) and $z = \frac{1}{2}$, and finally arrives at z_{j+1} (again see Fig. 9.8). This means that

$$\gamma_2^* : \{g_1, g_2, g_3, g_4\} \to \{g_1, g_2, g_3, g_4\}, \qquad \gamma_2^*(g_j) = g_{j+1}.$$

This computation also shows that

$$\gamma_2^*(\gamma_2^*) = (\gamma_2 \cdot \gamma_2)^* : \{g_1, g_2, g_3, g_4\} \to \{g_1, g_2, g_3, g_4\}, \quad \gamma_2^*(\gamma_2^*)g_j = g_{j+2}.$$

Since there are only two basic homotopy classes in L_B, namely those curves that do not loop around the critical value $z = \frac{1}{16}$ (for example, γ_1) and those that loop around $z = \frac{1}{16}$ (for example, $\gamma_2, \gamma_2 \cdot \gamma_2$, and $\gamma_2 \cdot \gamma_2 \cdot \gamma_2$), we see that the monodromy group $\mathscr{G}_B$ is a cyclic group of order four that is isomorphic to the subgroup of S_4 generated by the 4-cycle (1 2 3 4). This last observation will be important later.

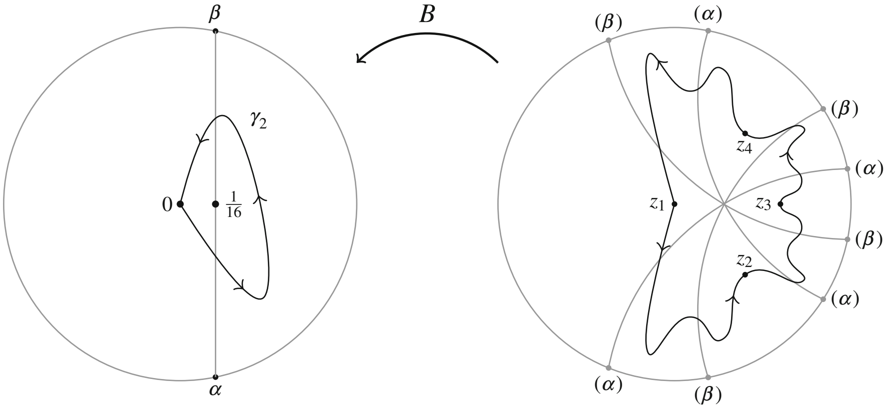


Fig. 9.8 The curve γ_2 (*left*) and the inverse image curves for γ_2 (*right*). Notice how γ_2 first crosses δ between the critical value $\frac{1}{16}$ and α, turns, recrosses δ between β and the critical value, before it returns to the origin. With this in mind, observe how $B^{-1}(\gamma_2)$ starts off at z_j, crosses $B^{-1}(\delta)$ between (α) and the critical point $z = \frac{1}{2}$, crosses $B^{-1}(\delta)$ between (β) and $z = \frac{1}{2}$, finally arriving at z_{j+1}

Example 9.4.2 The finite Blaschke product

$$b(z) = z^2\left(\frac{\frac{1}{2} - z}{1 - \frac{1}{2}z}\right)^2$$

is not in normalized form since it has zeros of order two at the origin and at $z = \frac{1}{2}$. One can normalize b to obtain a finite Blaschke product B that has four distinct zeros z_1, z_2, z_3, z_4 (with $z_1 = 0$), three critical points p_1, p_2, p_3, and two distinct critical values v_1, v_2. We are intentionally vague about the exact numbers involved since they are typographically cumbersome and were only selected to make the illustrations reasonable to view.

We connect the two critical values v_1 and v_2 with a curve δ that meets the circle at two points α and β. Note that this curve does not pass through the origin; see Fig. 9.9. As with the previous example, we draw the inverse image curves $B^{-1}(\delta)$. Observe that δ divides $\mathbb{D}$ into two regions, one of which contains the origin, while $B^{-1}(\delta)$ divides $\mathbb{D}$ into eight regions, four of which contain a zero of B.

Next we label the zeros of B counterclockwise as z_1, z_2, z_3, z_4 (with $z_1 = 0$) and, as usual, denote the branches of B^{-1} at 0 by g_1, g_2, g_3, g_4. For the four most basic types of homotopy classes in L_B, $\gamma_1, \gamma_2, \gamma_3, \gamma_4$, we draw $B^{-1}(\gamma_j)$ and take note of the permutations of the zeros; see Figs. 9.10, 9.11, 9.12, and 9.13.

Observe from the corresponding drawings, and the discussion from the previous example, that

$$\gamma_1^* g_j = g_j, \qquad j = 1, 2, 3, 4,$$

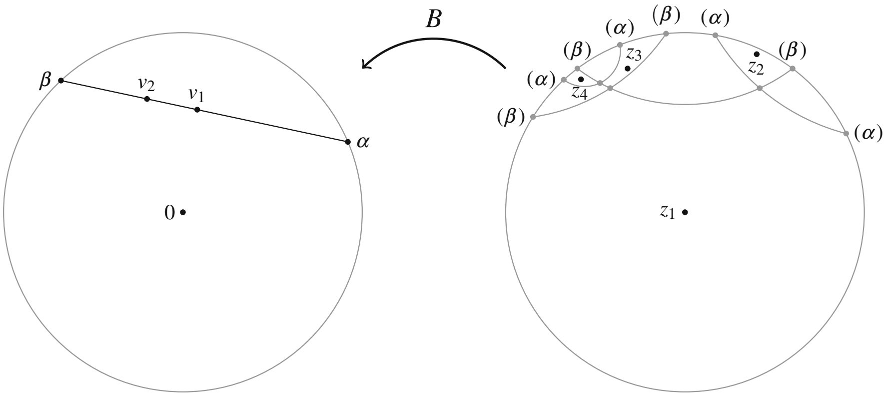


Fig. 9.9 The curve δ passing through the critical values (*left*) and the curves $B^{-1}(\delta)$ passing through the critical points (interior intersection points) (*right*)

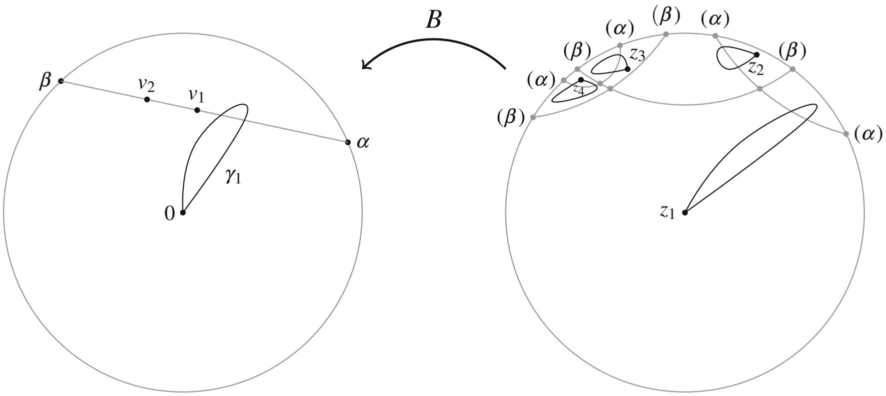


Fig. 9.10 The curve γ_1 (*left*) and $B^{-1}(\gamma_1)$ (*right*)

$$\gamma_2^* g_1 = g_2, \qquad \gamma_2^* g_3 = g_4,$$

$$\gamma_3^* g_1 = g_2, \qquad \gamma_3^* g_2 = g_3, \qquad \gamma_3^* g_3 = g_4, \qquad \gamma_3^* g_4 = g_1,$$

$$\gamma_4^* g_1 = g_3, \qquad \gamma_4^* g_2 = g_2, \qquad \gamma_4^* g_4 = g_4.$$

Equating g_j with j, we see that $\mathscr{G}_B$ is the subgroup of S_4 generated by the permutations

$$(1\ 2)(3\ 4), \quad (1\ 2\ 3\ 4), \quad (1\ 3).$$

In fact, $\mathscr{G}_B$ is isomorphic to the dihedral group of a square; see Exercise 9.3.

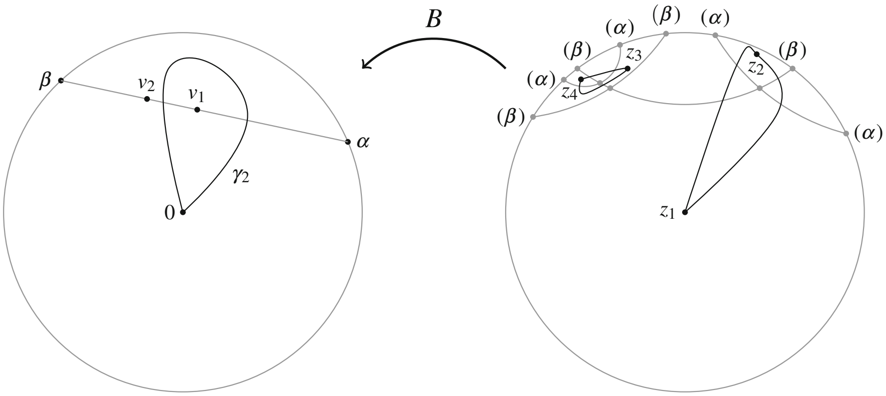


Fig. 9.11 The curve γ_2 (*left*) and $B^{-1}(\gamma_2)$ (*right*)

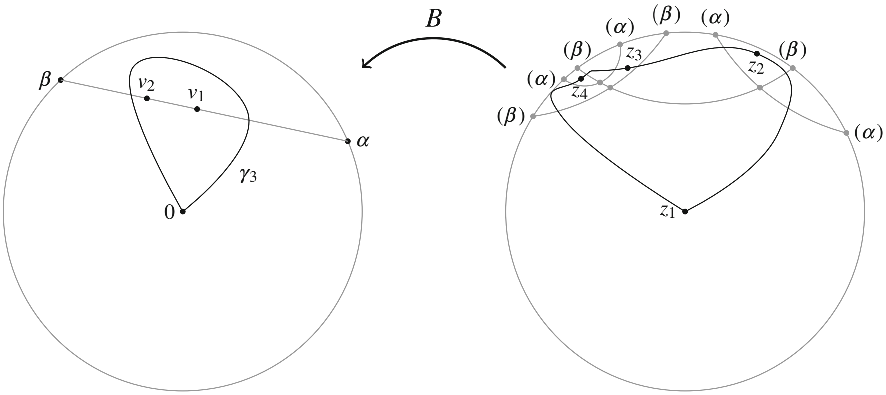


Fig. 9.12 The curve γ_3 (*left*) and $B^{-1}(\gamma_3)$ (*right*)

Example 9.4.3 The finite Blaschke product

$$b(z) = z^2 \cdot \frac{z - \frac{1}{3}}{1 - \frac{1}{3}z} \cdot \frac{z - \frac{1}{2}}{1 - \frac{1}{2}z}$$

is not in normalized form since it has a double zero at the origin. One can normalize b and obtain a normalized finite Blaschke product B. This normalized finite Blaschke product has four distinct zeros z_1, z_2, z_3, z_4 (with $z_1 = 0$), three distinct critical points c_1, c_2, c_3, and three distinct critical values v_1, v_2, v_3. In Fig. 9.14, we plot the curve δ that passes through the three critical values (along with α and β) together with the inverse image curves $B^{-1}(\delta)$ that intersect at the critical points.

From Figs. 9.15 and 9.16 we see that $\mathcal{G}_B$ contains a group element that transposes g_3 and g_4 and a group element that implements the permutation $g_1 \mapsto g_2 \mapsto g_3 \mapsto$

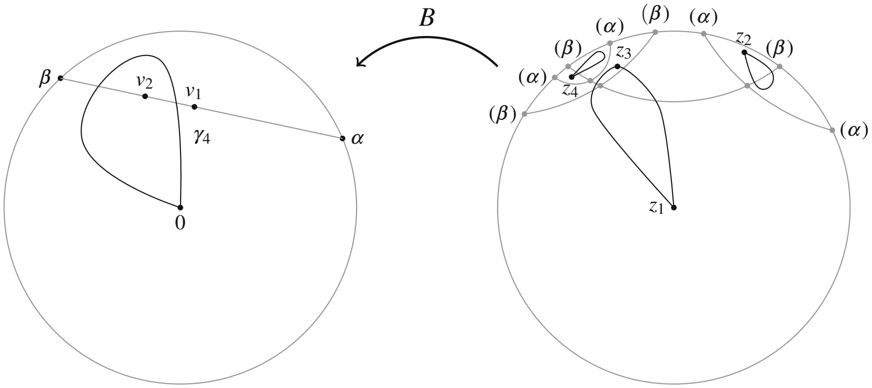


Fig. 9.13 The curve γ_4 (*left*) and $B^{-1}(\gamma_4)$ (*right*)

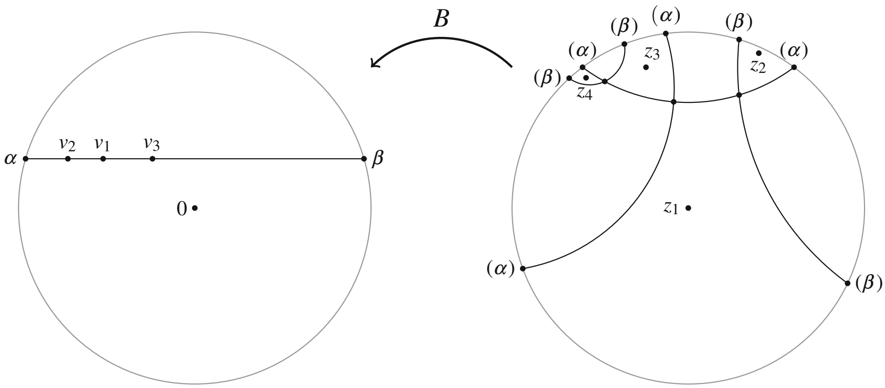


Fig. 9.14 The curve δ passing through the critical values (*left*) and the curves $B^{-1}(\delta)$ passing through the critical points (*right*)

g_4. In other words, $\mathscr{G}_B$ is isomorphic to the subgroup of S_4 that contains the cycles (3 4) and (1 2 3 4). Since these two cycles generate S_4, we conclude that $\mathscr{G}_B$ is isomorphic to S_4.

9.5 Primitive Versus Imprimitive

Let B denote a normalized finite Blaschke product of degree n with monodromy group $\mathscr{G}_B$. By the construction in the previous section, each $\gamma^* \in \mathscr{G}_B$ is a permutation of the branches $\{g_1, g_2, \ldots, g_n\}$ at 0. Also recall that $\mathbf{e}$ is the trivial loop and hence $\mathbf{e}^*$ is the identity element of $\mathscr{G}_B$. The group $\mathscr{G}_B$ acts on $\{g_1, g_2, \ldots, g_n\}$ in that the function

$$\phi : \mathscr{G}_B \times \{g_1, g_2, \ldots, g_n\} \to \{g_1, g_2, \ldots, g_n\}, \qquad \phi(\gamma^*, g_j) = \gamma^* g_j,$$

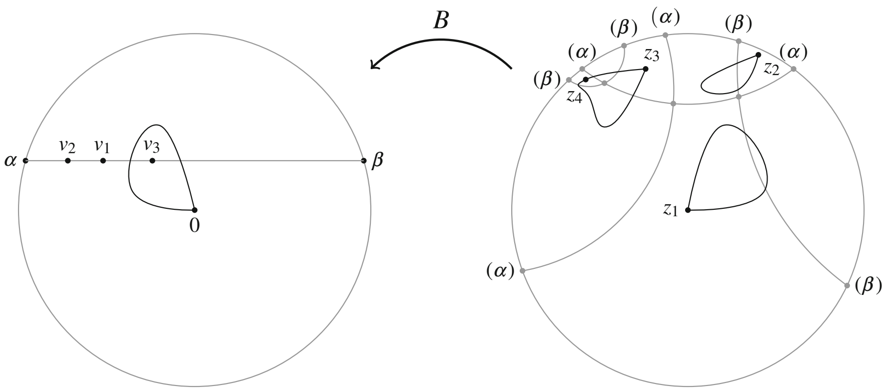


Fig. 9.15 A curve circulating around v_3 (*left*) and its corresponding pre-image (*right*). Observe how this yields the identity on g_1 and g_2 and reverses g_3 and g_4

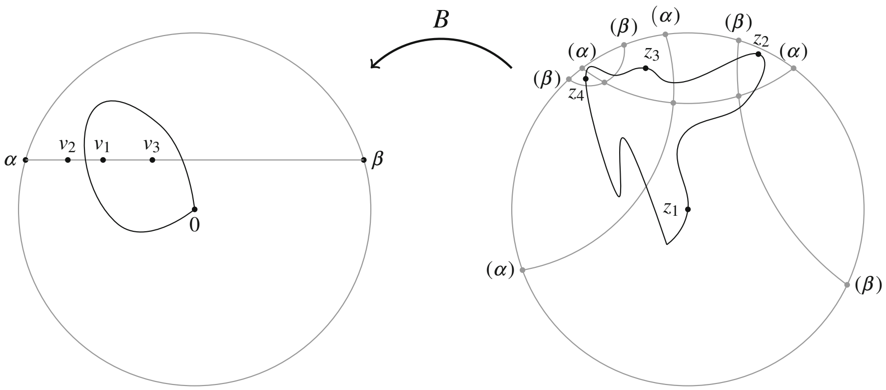


Fig. 9.16 A curve circulating around v_1 and v_3 and its corresponding pre-image. Notice how this yields the cycle $g_1 \mapsto g_2 \mapsto g_3 \mapsto g_4$

satisfies

$$\mathbf{e}^* g_j = g_j, \qquad 1 \leqslant j \leqslant n, \qquad \text{(identity axiom)},$$

and

$$(\gamma^* \delta^*) g_j = \gamma^* (\delta^* g_j), \qquad 1 \leqslant j \leqslant n, \qquad \text{(compatibility axiom)}.$$

Definition 9.5.1 For a finite group G acting on a finite set X, we say that G *respects a partition* $\mathscr{P}$ *of* X if $gP \in \mathscr{P}$ for all $P \in \mathscr{P}$ and all $g \in G$.

Here is another way of thinking about this concept that yields a little more information.

Lemma 9.5.2 *For a finite group G acting on a finite set X, the following are equivalent.*

(a) G respects a partition $\mathscr{P}$ of X.
(b) For each $g \in G$ and $P \in \mathscr{P}$, there is a $P' \in \mathscr{P}$ such that $gP \subseteq P'$.

Proof The proof of (a) $\implies$ (b) is automatic. To see that (b) $\implies$ (a), observe that if $gP \subseteq P'$ for all $g \in G$ and $P \in \mathscr{P}$, then $g^{-1}P' \subseteq P''$ for some $P'' \in \mathscr{P}$. Note that $P \subseteq g^{-1}P'$ and since $\mathscr{P}$ is a partition of X, we have $P = g^{-1}P' = P''$. Hence, $gP = P'$. □

Any partition $\mathscr{P}$ that G respects partitions X into subsets of equal size. The size of these sets is called the *order of the partition*. There are at least two partitions of X that G respects, namely $\{X\}$, the whole set, and $\{\{x\} : x \in X\}$, the set of singletons.

Definition 9.5.3 If $\{X\}$ and $\{\{x\} : x \in X\}$ are the only two partitions that G respects, then the action of G on X is *primitive*. If there is another partition $\mathscr{P}$ of X that G respects, then the action of G on X is *imprimitive*.

Definition 9.5.4 The action of a group G on a set X is *transitive* if $Gx = X$ for some (and hence all) $x \in X$.

We will make use of the following classification of primitive group actions (see [81] or [135] for a proof).

Theorem 9.5.5 *Suppose that G is a group that acts transitively on a set X. Then G acts primitively on X if and only if for each $x \in X$, the stabilizer*

$$\{g \in G : gx = x\}$$

is a maximal subgroup of G. That is, there is no subgroup H of G such that $\{g \in G : gx = x\} \subsetneq H \subsetneq G$.

To apply this result to $G = \mathscr{G}_B$, we need the following result.

Theorem 9.5.6 *For a normalized finite Blaschke product B, the monodromy group $\mathscr{G}_B$ acts transitively on the branches $\{g_1, g_2, \ldots, g_n\}$ at 0.*

Proof Let δ be a curve that meets $\mathbb{T}$ at two distinct points, does not pass through the origin, and does not pass through any of the critical values of B. Let γ be a loop that starts at 0, circulates counterclockwise, surrounds all of the critical values of B, and meets δ in exactly two places, before it returns to zero; see Fig. 9.17. We leave it as an exercise (Exercise 9.1) to use the analysis from our previous examples to see that the powers of γ^* form a cycle of the branches $g_1, g_2, \ldots, g_n$ in the sense that

$$\gamma^* g_1 = g_2, \quad \gamma^{*2} g_2 = g_3 \quad \ldots \quad \gamma^{*(n-1)} g_{n-1} = g_n, \quad \gamma^{*n} g_n = g_1.$$

This proves that $\mathscr{G}_B$ is transitive. □

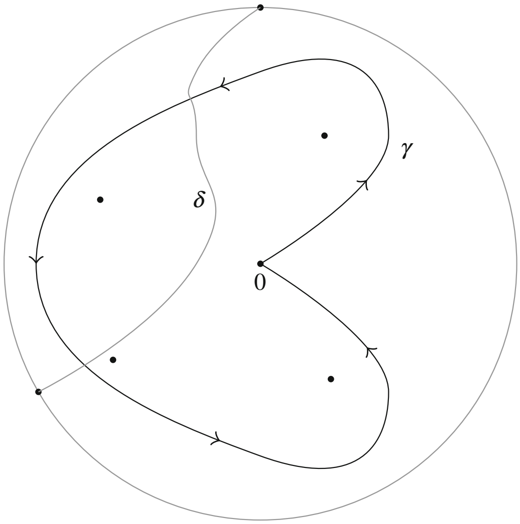


Fig. 9.17 The curves γ and δ from Theorem 9.5.6

9.6 Ritt's Theorem

With all of the pieces in place, we are ready for Ritt's theorem [119].

Theorem 9.6.1 (Ritt) *A normalized Blaschke product* B *of degree* $n > 1$ *is decomposable if and only if the monodromy group* $\mathscr{G}_B$ *acts imprimitively on the branches* $\{g_1, g_2, \ldots, g_n\}$ *of* B^{-1} *at* $z = 0$.

Proof We follow Cowen's proof from [28]. Suppose that $\mathscr{G}_B$ acts imprimitively on the branches $\{g_1, g_2, \ldots, g_n\}$ of B^{-1} at 0. Then there is a partition

$$\mathscr{P} = \{P_1, P_2, \ldots, P_m\}$$

of the branches $\{g_1, g_2, \ldots, g_n\}$ where, by the remark following Definition 9.5.1, each P_j has the same order k and hence $n = mk$. Renumbering the branches if necessary, we write

$$\begin{aligned} P_1 &= \{g_1, g_2, \ldots, g_k\}, \\ P_2 &= \{g_{k+1}, g_{k_2}, \ldots, g_{2k}\}, \\ &\vdots \\ P_m &= \{g_{(m-1)k+1}, g_{(m-1)k+2}, \ldots, g_{mk}\}. \end{aligned}$$

Recall the critical values $\mathscr{S}_B$ from (9.3.1) and the set $\widetilde{\mathscr{S}_B} = B^{-1}(\mathscr{S}_B)$ from (9.3.2). Each $g_j \circ B$ is arbitrarily continuable to $\mathbb{D}\backslash\widetilde{\mathscr{S}_B}$ since $B(\mathbb{D}\backslash\widetilde{\mathscr{S}_B}) \subseteq \mathbb{D}\backslash\mathscr{S}_B$ and g_j is arbitrarily continuable to $\mathbb{D}\backslash\mathscr{S}_B$. For z in some neighborhood of zero contained in $\mathbb{D}\backslash\widetilde{\mathscr{S}_B}$, define

$$D(z) = z \prod_{j=2}^{k} \frac{\overline{g_j(0)}}{|g_j(0)|} g_j(B(z))$$
$$= g_1(B(z)) \prod_{j=2}^{k} \frac{\overline{g_j(0)}}{|g_j(0)|} g_j(B(z)),$$

in which the last equality follows from the identity $g_1(B(z)) = z$, which in turn follows from the fact that $B(g_1(B(z))) = B(z)$ and $g_1(B(0)) = 0$. Since each factor $g_j \circ B$ is arbitrarily continuable on $\mathbb{D}\backslash\widetilde{\mathscr{S}_B}$ so will D be with the appropriate product formula holding for all $z \in \mathbb{D}\backslash\widetilde{\mathscr{S}_B}$.

Let δ be a closed loop in $\mathbb{D}\backslash\widetilde{\mathscr{S}_B}$ that includes 0. Continuing D along δ is essentially the same as continuing the function

$$g_1(w) \prod_{j=2}^{k} \frac{\overline{g_j(0)}}{|g_j(0)|} g_j(w)$$

along the closed curve $\gamma = B \circ \delta$ in $\mathbb{D}\backslash\mathscr{S}_B$. We claim that $\gamma^*(g_1) = g_1$. Indeed, $g_1(B(z)) = z$ is single valued on $\mathbb{C}$ and hence continuing g_1 along γ is the same as continuing z along δ. Thus, $\gamma^* g_1 = g_1$. Now observe that since $\mathscr{G}_B$ respects the partition $\mathscr{P}$, we have $\gamma^* P_1 = P_s$ for some $1 \leqslant s \leqslant m$. However, we already know that $\gamma^* g_1 = g_1$ and so it must be the case that $\gamma^* P_1 = P_1$. All of this implies that continuing D along δ results only in rearranging the factors in D and hence the continuation of D is D. Since D is arbitrarily continuable in $\mathbb{D}\backslash\widetilde{\mathscr{S}_B}$ and single valued in a neighborhood of 0, we conclude that D is single valued on all of $\mathbb{D}\backslash\widetilde{\mathscr{S}_B}$. But since we also have $|D| < 1$ on $\mathbb{D}\backslash\widetilde{\mathscr{S}_B}$ and $\widetilde{\mathscr{S}_B}$ is a finite set, D defines an analytic function on all of $\mathbb{D}$.

Everything we have done so far can also be done not only on $\mathbb{D}$ but also in an open neighborhood of $\mathbb{D}^-$ where B defines an analytic function (just avoid the poles of B). Thus, D defines an analytic function in an open neighborhood of $\mathbb{D}^-$. For $\xi \in \mathbb{T}$, the $g_j(B(\xi))$ is unimodular and hence D is unimodular on $\mathbb{T}$. By Fatou's theorem (Theorem 3.5.2), D is a finite Blaschke product. In particular, D is a finite Blaschke product whose zeros are $g_1(0), g_2(0), \ldots, g_k(0)$ and hence D has degree equal to k, the order of the partition $\mathscr{P}$. From the formula defining D, we have $D'(0) > 0$.

We now claim that

$$D(g_1(0)) = D(g_2(0)) = \cdots = D(g_k(0)),$$
$$D(g_{k+1}(0)) = D(g_{k+2}(0)) = \cdots = D(g_{2k}),$$
$$D(g_{2k+1}(0)) = D(g_{2k+2}(0)) = \cdots = D(g_{3k}(0)),$$
$$\vdots$$
$$D(g_{(m-1)k+1}(0))) = \cdots = D(g_{mk}(0)).$$

We will do this by showing that

$$D(g_{rk+j}(0)) = \left(\prod_{l=2}^{k} \frac{\overline{g_l(0)}}{|g_l(0)|}\right)\left(\prod_{l=1}^{k} g_{rk+l}(0)\right).$$

Observe that the right-hand side of the preceding equation is independent of j. To prove this formula, let γ be a closed curve in $\mathbb{D}\backslash\mathscr{S}_B$ such that

$$\gamma^* g_1 = g_{rk+j}.$$

Note how we are using the transitivity of $\mathscr{G}_B$ (Theorem 9.5.6). Let δ be the lift of γ to $\mathbb{D}\backslash\widetilde{\mathscr{S}_B}$ (via B^{-1}) with $\delta(0) = 0$. Thus, $\delta(1) = g_{rk+j}(0)$. By definition, $D(g_{rk+j}(0))$ is the continuation of

$$z \prod_{l=2}^{k} \frac{\overline{g_l(0)}}{|g_l(0)|} g_l(B(z))$$

along δ. Since $\gamma^* g_1 = g_{rk+j}$ and $\mathscr{G}_B$ respects the partition $\mathscr{G}_B$ we see that $\gamma^* P_1 = P_r$ and hence

$$D(g_{rk+j}(0)) = \left(\prod_{l=2}^{k} \frac{\overline{g_l(0)}}{|g_l(0)|}\right)\left(\prod_{l=1}^{k} g_{rk+l}(0)\right). \tag{9.6.2}$$

Let C be the finite Blaschke product with $C'(0) > 0$ and whose zeros are

$$0 = D(g_1(0)), \quad D(g_{k+1}(0)), \quad D(g_{2k+1}(0)), \ldots, \quad D(g_{(m-1)k+1}(0)).$$

By (9.6.2), $C \circ D$ is a Blaschke product with $(C \circ D)'(0) > 0$ and with zeros

$$g_1(0), \; g_2(0), \ldots, \; g_n(0).$$

Since B is a finite Blaschke product with the same zeros and $B'(0) > 0$, we conclude that $B = C \circ D$; see Exercise 9.4.

To prove the converse, suppose that $B = C \circ D$ (where C and D are finite Blaschke products of degree greater than one) and $\{g_1, g_2, \ldots, g_n\}$ are the branches of B^{-1} at zero. For each g_j, we know that $D \circ g_j$ is a branch of C^{-1} at zero. This allows us to define an equivalence relation on $\{g_1, g_2, \ldots, g_n\}$ by

$$g_j \sim g_{j'} \quad \Longleftrightarrow \quad D \circ g_j = D \circ g_{j'}$$

on their common domain. This equivalence relation produces a partition $\mathscr{P}$ of $\{g_1, g_2, \ldots, g_n\}$. That $\mathscr{G}_B$ respects $\mathscr{P}$ is a consequence of the permanence of functional relations [100]. Indeed, observe that for all curves $\gamma \in L_B$ we have

$$D \circ g_j = D \circ g_{j'} \implies D \circ (\gamma^* g_j) = D \circ (\gamma^* g_{j'}).$$

In other words, $\mathscr{G}_B$ respects the partition $\mathscr{P}$. □

9.7 Examples of Decomposability

Example 9.7.1 Revisiting Example 9.4.1, where B is a normalization of $b(z) = z^4$, we saw that $\mathscr{G}_B$ is a cyclic group of order 4 that can be viewed as the cyclic group generated by the 4-cycle (1 2 3 4) acting on $\{1, 2, 3, 4\}$. Here we identify 1 with the zero z_1, 2 with the zero z_2, and so forth. A quick verification confirms that $\mathscr{G}_B$ acts transitively. One can also see that the stabilizer of $\{1\}$ is the trivial group, which is properly contained in the proper subgroup H generated by the 2-cycle $(1, 3)$. Thus, $\mathscr{G}_B$ acts imprimitively on $\{g_1, g_2, g_3, g_4\}$ which makes B (and hence b) decomposable. Indeed, $b = z^2 \circ z^2$.

Example 9.7.2 Revisiting Example 9.4.2, where B is a normalization of

$$b(z) = z^2\left(\frac{\frac{1}{2} - z}{1 - \frac{1}{2}z}\right)^2,$$

we saw that $\mathscr{G}_B$ was isomorphic to the dihedral group D_4, which acts transitively. One can view D_4 as acting on the vertices of a square labeled in order as $1, 2, 3, 4$. The elements of D_4 (using cycle notation) are

$$(1\ 2\ 3\ 4),\quad (1\ 3),\quad (1\ 4\ 3\ 2),\quad (1\ 3),\quad (2\ 4),\quad (1\ 2)(3\ 4),\quad (1\ 4)(2\ 3),\quad (1).$$

Direct computation reveals that the stabilizer of $\{1\}$ is the subgroup $\langle(2\ 4)\rangle$ (the subgroup generated by (2 4)) and that

$$\langle(2\ 4)\rangle \subsetneq \langle(2\ 4), (1\ 3)\rangle \subsetneq D_4.$$

Thus, $\mathscr{G}_B$ acts imprimitively on $\{g_1, g_2, g_3, g_4\}$ which makes B (and hence b) decomposable. Indeed,

$$b = z^2 \circ \left(z\frac{\frac{1}{2} - z}{1 - \frac{1}{2}z}\right).$$

Example 9.7.3 Revisiting Example 9.4.3, where B is a normalization of

$$b(z) = z^2 \cdot \frac{z - \frac{1}{3}}{1 - \frac{1}{3}z} \cdot \frac{z - \frac{1}{2}}{1 - \frac{1}{2}z},$$

we saw that $\mathscr{G}_B$ was isomorphic to S_4, which acts transitively. The stabilizer of $\{1\}$ is isomorphic to the group of permutations of $\{2, 3, 4\}$, a group of order 6. A proper subgroup of S_4 that contains the stabilizer would have to be of order 12 by Lagrange's Theorem. However, the only subgroup of S_4 of order 12 is A_4, the alternating group on four letters. However, A_4 does not contain the 2-cycle $(2\ 3)$, which belongs to the stabilizer of $\{1\}$, since this element is a transposition and hence is not an even permutation. Thus, the stabilizer of $\{1\}$ is a maximal subgroup of S_4. Analogous arguments show that the stabilizers of $\{2\}$, $\{3\}$, and $\{4\}$ are also maximal subgroups of S_4. Consequently, $\mathscr{G}_B$ acts primitively on $\{g_1, g_2, g_3, g_4\}$ and hence B and b are indecomposable finite Blaschke products.

9.8 Notes

Ritt's Theorem for Polynomials

There is a statement of Ritt's theorem for polynomials. A polynomial p can be written as a nontrivial composition $p = q \circ r$ of polynomials q and r if and only if the associated monodromy group is imprimitive [119] (see also [108]).

Ritt's Theorem Redux

There are several other characterizations of decomposable finite Blaschke products [23, 30, 130].

Ritt's Theorem and Cyclic Groups

For a finite Blaschke product B of degree n, we defined the group

$$G_B = \{u \in \mathscr{C} : B \circ u = B\}$$

and proved in Theorem 9.1.6 that this group was cyclic of order n. It turns out that we can rephrase Ritt's theorem in terms of G_B. From [30] we have the following theorem: a finite Blaschke product B of degree $n = mk$, in which $m > 1$, is a composition of two nontrivial finite Blaschke products $B = C \circ D$ if and only if there is a finite Blaschke product D of degree $k > 1$ such that the group G_D is generated by u^m, where u is a generator of the group G_B.

Chebyshev Blaschke Products

In [108] there is an exposition of *Chebyshev Blaschke products*. The Chebyshev polynomials

$$T_n(z) = \cos(n \arccos z)$$

are well-known orthogonal polynomials and have the nesting property

$$T_{mn} = T_m \circ T_n$$

with respect to function composition. The monodromy group of T_n was computed by Ritt [119]. There is also a family of Chebyshev Blaschke products $B_{n,\tau}$, where $n \in \mathbb{N}$ and $\tau \in i\mathbb{R}_+$. These Blaschke products satisfy the nesting property

$$B_{mn,\tau} = B_{m,n\tau} \circ B_{n,\tau}$$

and one can compute the monodromy group of $B_{n,\tau}$.

An Interesting Approximation Result

The indecomposable finite Blaschke products are uniformly dense in the set of all finite Blaschke products [22] and thus, by Carathéodory's Theorem (Theorem 4.1.1), such indecomposable finite Blaschke products are dense in the unit ball of H^∞.

Further Examples

Further, more complicated, examples of monodromy groups associated with finite Blaschke products were worked out in a thesis of B. Sokolowsky [130].

9.9 Exercises

9.1 Maintaining the notation of Theorem 9.5.6 and its proof, show that the powers of γ^* form a cycle of the branches in the sense that

$$\gamma^* g_1 = g_2, \qquad \gamma^{*2} g_2 = g_3, \quad \ldots \quad \gamma^{*(n-1)} g_{n-1} = g_n, \qquad \gamma^{*n} g_n = g_1.$$

9.2 Show that if a finite group G acts on a finite set X, then the cardinality of X divides the cardinality of G.
Hint: Use the orbit-stabilizer theorem.

9.3 Finish the details of the proof of Example 9.4.2 and show that $\mathscr{G}_B$ is isomorphic to the dihedral group of a square.

9.4 Show that if B_1 and B_2 are normalized finite Blaschke products with the same zeros, then $B_1 \equiv B_2$.

9.5 Which subgroups of S_3 are realizable as monodromy groups of finite Blaschke products? Which subgroups of S_4?

9.6 Suppose that B is a finite Blaschke product of degree n. If B has more than $\frac{n}{2}$ critical values, then B is decomposable [130].

Chapter 10
Finite Blaschke Products and Operator Theory

In this chapter we explore some of the connections that finite Blaschke products make with operators on Hilbert spaces. In particular, we focus on norms of contractions and the mapping properties of the numerical range. A review of some relevant operator theory notions such as the norm, spectrum, functional calculus, and spectral mapping theorem can be found in Appendix A.6.

10.1 Contractions

A function of the form

$$p(\zeta) = \sum_{n=-N}^{N} a_n \zeta^n, \qquad \zeta \in \mathbb{T}, \tag{10.1.1}$$

is a *trigonometric polynomial*. Although initially defined on $\mathbb{T}$, every trigonometric polynomial is defined and analytic on $\mathbb{C}\backslash\{0\}$. The term "trigonometric" stems from the fact that if we write $\zeta = e^{i\theta} = \cos\theta + i\sin\theta$ and substitute this into (10.1.1), the result is a complex linear combination of sines and cosines. If the trigonometric polynomial p has no negatively indexed coefficients, that is,

$$p(\zeta) = \sum_{n=0}^{N} a_n \zeta^n, \qquad \zeta \in \mathbb{T},$$

then p is an *analytic polynomial*. This terminology reflects the fact that $p(z) = \sum_{n=0}^{N} a_n z^n$ is a polynomial in the usual sense. The term "polynomial," when used without modifiers, refers to an analytic polynomial.

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_10

Each trigonometric polynomial is continuous on $\mathbb{T}$ and hence is bounded there. Consequently, we may define

$$\|p\|_\infty = \max_{\zeta\in\mathbb{T}} |p(\zeta)|.$$

If p is an analytic polynomial, then the Maximum Modulus Principle implies that

$$\max_{z\in\mathbb{D}} |p(z)| = \max_{\zeta\in\mathbb{T}} |p(\zeta)|.$$

A trigonometric polynomial p is *positive* if $p(\zeta) \geqslant 0$ for all $\zeta \in \mathbb{T}$. If q is a trigonometric polynomial, then

$$p = \overline{q}q = |q|^2$$

is a positive trigonometric polynomial. The following theorem of Fejér [47] and Riesz [118] asserts that every positive trigonometric polynomial arises in this manner. Many generalizations of this result are discussed in [35].

Theorem 10.1.2 (Fejér–Riesz) *If p is a positive trigonometric polynomial, then $p = |q|^2$ on $\mathbb{T}$ for some analytic polynomial q with no roots in $\mathbb{D}$.*

Proof Suppose that $p(\zeta) = \sum_{n=-N}^{N} a_n \zeta^n$ is a positive trigonometric polynomial. Then $a_N = a_{-N}$ since $p = \overline{p}$ on $\mathbb{T}$ and hence we may assume that $a_{-N} \neq 0$. Thus, $f(z) = z^N p(z)$ is an analytic polynomial of degree $2N$ with $f(0) \neq 0$. In particular, the zeros in $\mathbb{C}$ of $f(z)$ are precisely the zeros of $p(z)$. Since

$$p(\zeta) = \overline{p(\zeta)} = \overline{p(1/\overline{\zeta})}, \qquad \zeta \in \mathbb{T},$$

we conclude that $p(z) = \overline{p(1/\overline{z})}$ for $z \in \mathbb{C}\backslash\{0\}$. Consequently, each zero $\alpha \neq 0$ of f in $\mathbb{C}\backslash\mathbb{T}$ occurs as a pair $\{\alpha, 1/\overline{\alpha}\}$ with matching multiplicities.

Suppose that $\beta \in \mathbb{T}$ is a zero of f of order m and fix a small neighborhood U of β whose closure includes no other zeros of f. For $\epsilon > 0$, we observe that $p + \epsilon$ is a trigonometric polynomial that is strictly positive on $\mathbb{T}$. In particular, it has no zeros on $\mathbb{T}$. Since $p + \epsilon$ converges uniformly to p on $\mathbb{C}$ as $\epsilon \to 0$, it follows that $f_\epsilon(z) = z^N(p(z) + \epsilon)$ converges uniformly to f on U as $\epsilon \to 0$. Hurwitz' theorem [26, p. 152] says that for sufficiently small $\epsilon > 0$, f_ϵ has exactly m zeros in U. The reasoning in the first paragraph ensures that the zeros of f_ϵ occur in pairs $\{\alpha, 1/\overline{\alpha}\}$ with matching multiplicities. Moreover, f_ϵ has no zeros on $\mathbb{T}$ since $p + \epsilon$ does not. We conclude from this that m is even.

The preceding discussion implies that

$$f(z) = c\prod_{i=1}^{N}(z - z_i)(1/\overline{z_i} - z),$$

in which $z_1, z_2, \ldots, z_N$ are in $\mathbb{C}\backslash\mathbb{D}$ and $c \neq 0$. Then

$$\begin{aligned} p(\zeta) &= \zeta^{-N} f(\zeta) \\ &= c\overline{\zeta}^N \prod_{i=1}^{N} (\zeta - z_i)(1/\overline{z_i} - \zeta) \\ &= C \prod_{i=1}^{N} (\zeta - z_i)(\overline{\zeta} - \overline{z_i}) \\ &= C \prod_{i=1}^{N} (\zeta - z_i) \prod_{i=1}^{N} \overline{(\zeta - z_i)}, \end{aligned}$$

in which

$$C = c \prod_{j=1}^{n} \frac{1}{\overline{z_j}} \neq 0.$$

Since p is a positive trigonometric polynomial, we conclude that $C > 0$. Thus, $p = q\overline{q} = |q|^2$, in which

$$q(\zeta) = \sqrt{C} \prod_{i=1}^{N} (\zeta - z_i)$$

is an analytic polynomial with no roots in $\mathbb{D}$. □

Let $\mathscr{H}$ denote a Hilbert space and let $\mathscr{L}(\mathscr{H})$ denote the set of all bounded linear operators on $\mathscr{H}$. If p is given by (10.1.1), we may write

$$p(\zeta) = a_0 + \sum_{n=1}^{N} a_n \zeta^n + \sum_{n=1}^{N} a_{-n} \overline{\zeta^n} \tag{10.1.3}$$

and define the operator

$$p(T) = a_0 I + \sum_{n=1}^{N} a_n T^n + \sum_{n=1}^{N} a_{-n} T^{*n} \tag{10.1.4}$$

for $T \in \mathscr{L}(\mathscr{H})$. The map $p \mapsto p(T)$ is well defined and linear; see Exercise 10.1. Less obvious is the fact that this map preserves positivity. This is Lemma 10.1.10 below: if $T \in \mathscr{L}(\mathscr{H})$ is a contraction and p is a positive trigonometric polynomial, then $p(T) \geqslant 0$ (recall the definition of a positive operator from Appendix A.6). Our proof relies upon a powerful result of Béla Szőkefalvi-Nagy which asserts that, for many purposes, a contraction can be replaced by a unitary operator [131]. The unitary operator U constructed in the following theorem is called a *unitary dilation* of T.

Theorem 10.1.5 (Szőkefalvi-Nagy Dilation Theorem) *If $T \in \mathscr{L}(\mathscr{H})$ is a contraction, then there is a Hilbert space $\mathscr{K}$ that contains $\mathscr{H}$ and a unitary $U \in \mathscr{L}(\mathscr{K})$ such that*

$$T^n = P_{\mathscr{H}} U^n|_{\mathscr{H}}, \quad n = 0, 1, \ldots, \tag{10.1.6}$$

where $P_{\mathscr{H}}$ is the orthogonal projection of $\mathscr{K}$ onto $\mathscr{H}$.

Proof If T is a contraction, then $I - T^*T$ and $I - TT^*$ are positive and hence the *defect operators*

$$D_T = \sqrt{I - T^*T} \qquad \text{and} \qquad D_{T^*} = \sqrt{I - TT^*}$$

can be defined by the functional calculus for self-adjoint operators. Furthermore, we also have

$$D_{T^*}T = TD_T, \tag{10.1.7}$$

which implies that

$$D_{S^*}S = S^*D_{S^*} = 0$$

for any isometry S. The identity (10.1.7) requires a polynomial approximation argument along with the identity $p(T^*T)T = Tp(TT^*)$ for any analytic polynomial p; see Exercise 10.2.

Then

$$S = \begin{bmatrix} T & 0 & 0 & 0 \cdots \\ D_T & 0 & 0 & 0 \cdots \\ 0 & I & 0 & 0 \cdots \\ 0 & 0 & I & 0 \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{10.1.8}$$

which acts on the Hilbert space

$$\mathscr{J} = \bigoplus_{i=1}^{\infty} \mathscr{H},$$

is an isometric dilation of T. That is, $S^*S = I$ (which says that S is an isometry) and $p(T)$ equals the restriction of $p(S)$ to the first direct summand for each analytic polynomial p; see Exercise 10.4. Now observe that

$$U = \begin{bmatrix} S & D_{S^*} \\ 0 & -S^* \end{bmatrix}, \tag{10.1.9}$$

which acts on

$$\mathscr{K} = \mathscr{J} \oplus \mathscr{J},$$

is a unitary dilation of the isometry S; see Exercise 10.5. In other words, U is unitary and $p(S)$ equals the restriction of $p(U)$ to the first direct summand. If we identify the original $\mathscr{H}$ upon which T acts with the first direct summand of the first direct summand $\mathscr{J}$ in the decomposition of $\mathscr{K}$, then $p(T)$ equals the restriction of $p(U)$ to a subspace of $\mathscr{K}$. □

Lemma 10.1.10 *If $T \in \mathscr{L}(\mathscr{H})$ is a contraction and p is a positive trigonometric polynomial, then $p(T) \geqslant 0$.*

Proof The Fejér–Riesz theorem implies that $p = |q|^2$ on $\mathbb{T}$ for some analytic polynomial q. Let $U \in \mathscr{L}(\mathscr{K})$ be a unitary dilation of $T \in \mathscr{L}(\mathscr{H})$ (Theorem 10.1.5) and let $P_{\mathscr{H}}$ denote the orthogonal projection from $\mathscr{K}$ onto $\mathscr{H}$. Since U is unitary, we have $UU^* = U^*U = I$ and hence $p(U) = q(U)^*q(U)$. For each $\mathbf{x} \in \mathscr{H}$,

$$\begin{aligned} \langle p(T)\mathbf{x}, \mathbf{x}\rangle_{\mathscr{H}} &= \langle P_{\mathscr{H}} p(U)\mathbf{x}, \mathbf{x}\rangle_{\mathscr{H}} \\ &= \langle p(U)\mathbf{x}, \mathbf{x}\rangle_{\mathscr{K}} \\ &= \langle q(U)^* q(U)\mathbf{x}, \mathbf{x}\rangle_{\mathscr{K}} \\ &= \langle q(U)\mathbf{x}, q(U)\mathbf{x}\rangle_{\mathscr{K}} \\ &= \|q(U)\mathbf{x}\|_{\mathscr{K}}^2 \\ &\geqslant 0. \end{aligned}$$

Thus, $p(T) \geqslant 0$. □

An alternate proof of Lemma 10.1.10 is outlined in Exercise 10.3. We are now ready to prove a seminal result of von Neumann (see Exercise 10.6 for another proof).

Theorem 10.1.11 (von Neumann's Inequality [136]) *If T is a contraction and p is an analytic polynomial, then*

$$\|p(T)\| \leqslant \|p\|_\infty.$$

Proof Let $T \in \mathscr{L}(\mathscr{H})$ be a contraction with unitary dilation $U \in \mathscr{L}(\mathscr{K})$. Without loss of generality, suppose that $\|p\|_\infty = 1$ and consider the positive trigonometric polynomial

$$q = 1 - |p|^2 = 1 - \overline{p}p.$$

Then $q(U) = I - p(U)^* p(U)$ since $UU^* = U^*U = I$. For $\mathbf{x} \in \mathscr{K}$, Lemma 10.1.10 ensures that

$$\begin{aligned} 0 &\leqslant \langle q(U)\mathbf{x}, \mathbf{x}\rangle_{\mathscr{K}} \\ &= \langle (I - p(U)^* p(U))\mathbf{x}, \mathbf{x}\rangle_{\mathscr{K}} \\ &= \langle I\mathbf{x}, \mathbf{x}\rangle - \langle p(U)^* p(U)\mathbf{x}, \mathbf{x}\rangle_{\mathscr{K}} \\ &= \langle \mathbf{x}, \mathbf{x}\rangle - \langle p(U)\mathbf{x}, p(U)\mathbf{x}\rangle_{\mathscr{K}} \\ &= \|\mathbf{x}\|_{\mathscr{K}}^2 - \|p(U)\mathbf{x}\|_{\mathscr{K}}^2 \end{aligned}$$

and hence $\|p(U)\mathbf{x}\|_{\mathscr{K}} \leqslant \|\mathbf{x}\|_{\mathscr{K}}$ for all $\mathbf{x} \in \mathscr{K}$. Since $p(T) = P_{\mathscr{H}} p(U)|_{\mathscr{H}}$, we conclude that $\|p(T)\mathbf{x}\|_{\mathscr{H}} \leqslant \|\mathbf{x}\|_{\mathscr{H}}$ for $\mathbf{x} \in \mathscr{H}$. This completes the proof. □

The *Wiener algebra* $\mathscr{W}(\mathbb{D})$ consists of all analytic functions

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

on $\mathbb{D}$ such that

$$\sum_{n=0}^{\infty} |a_n| < \infty. \tag{10.1.12}$$

As its name suggests, the Wiener algebra is indeed an algebra; see Exercise 10.7. For each $f \in \mathscr{W}(\mathbb{D})$, the summability condition (10.1.12) guarantees that the Taylor polynomials

$$p_N(z) = \sum_{n=0}^{N} a_n z^n,$$

converge uniformly to f on $\mathbb{D}^-$. Consequently,

$$\|f\|_{\infty} = \sup_{\zeta \in \mathbb{T}} |f(\zeta)| = \lim_{N \to \infty} \sup_{\zeta \in \mathbb{T}} |p_N(\zeta)|. \tag{10.1.13}$$

The property (10.1.12) implies that for any contraction $T \in \mathscr{L}(\mathscr{H})$, the sequence $p_N(T)$ is Cauchy with respect to the operator norm. Thus, we can define the operator

$$f(T) = \sum_{n=0}^{\infty} a_n T^n = \lim_{N \to \infty} \sum_{n=0}^{N} a_n T^n. \tag{10.1.14}$$

Theorem 10.1.11 and (10.1.13) now yield

$$\|f(T)\| \leqslant \sup_{\zeta \in \mathbb{T}} |f(\zeta)|. \tag{10.1.15}$$

If $\text{Hol}(\mathbb{D}^-)$ denotes the set of functions that are analytic in a neighborhood of $\mathbb{D}^-$, then $\text{Hol}(\mathbb{D}^-) \subseteq \mathscr{W}(\mathbb{D})$. Indeed, for $f = \sum_{n=0}^{\infty} a_n z^n \in \text{Hol}(\mathbb{D}^-)$, the Cauchy–Hadamard formula for the radius of convergence of a power series tells us that

$$\limsup_{n\to\infty} |a_n|^{1/n} < 1.$$

Thus, we can define $f(T)$ for any $f \in \text{Hol}(\mathbb{D}^-)$ and any contraction T. Moreover, we have the operator norm estimate (10.1.15).

Corollary 10.1.16 *If T is a contraction and*

$$B(z) = \xi \prod_{k=1}^{n} \frac{z_k - z}{1 - \overline{z_k} z}$$

is a finite Blaschke product, then $B(T)$ is a contraction. Furthermore,

$$B(T) = \xi \prod_{k=1}^{n} (z_k I - T)(I - \overline{z_k} T)^{-1}. \tag{10.1.17}$$

Proof Observe that $B(T)$ is well defined since $B \in \text{Hol}(\mathbb{D}^-) \subseteq \mathscr{W}(\mathbb{D})$. It is a contraction by (10.1.15). The formula for $B(T)$ follows since $B(T)$, defined by the power series in (10.1.14), and the power series expansion of the right-hand side of (10.1.17), are the same. □

10.2 Norms of Contractions

From Corollary 10.1.16 we know that if T is a contraction and

$$B(z) = \xi \prod_{j=1}^{n} \frac{z_j - z}{1 - \overline{z_j} z} \tag{10.2.1}$$

is a finite Blaschke product of degree n, then $B(T)$ is also a contraction. When is $\|B(T)\| = 1$?

Theorem 10.2.2 (Gau–Wu [63]) *Suppose T is a contraction and $n \in \mathbb{N}$. The following are equivalent.*

(a) $\|B(T)\| = 1$ *for some finite Blaschke product of degree n.*
(b) $\|B(T)\| = 1$ *for every finite Blaschke product of degree n.*
(c) $\|T^n\| = 1$.

We closely follow the original proof from [63], which requires two lemmas. The first lemma is the asymptotic version of the fact that

$$\|T\mathbf{x}\| = \|\mathbf{x}\| \iff \mathbf{x} \in \ker(I - T^*T)$$

whenever T is a contraction.

Lemma 10.2.3 *If T is a contraction and $\mathbf{x}_k$ is a sequence such that $\|\mathbf{x}_k\| \to 1$, then*

$$\|T\mathbf{x}_k\| \to 1 \qquad \iff \qquad \|(1 - T^*T)\mathbf{x}_k\| \to 0.$$

Proof First observe that

$$\begin{aligned}
\|(I - T^*T)\mathbf{x}_k\|^2 &= \big\langle (I - T^*T)\mathbf{x}_k, (I - T^*T)\mathbf{x}_k \big\rangle \\
&= \langle \mathbf{x}_k, \mathbf{x}_k\rangle - \langle \mathbf{x}_k, T^*T\mathbf{x}_k\rangle - \langle T^*T\mathbf{x}_k, \mathbf{x}_k\rangle + \langle T^*T\mathbf{x}_k, T^*T\mathbf{x}_k\rangle \\
&= \|\mathbf{x}_k\|^2 - \big(\langle \mathbf{x}_k, T^*T\mathbf{x}_k\rangle + \overline{\langle \mathbf{x}_k, T^*T\mathbf{x}_k\rangle}\big) + \|T^*T\mathbf{x}_k\|^2 \\
&= \|\mathbf{x}_k\|^2 - 2\operatorname{Re}\langle \mathbf{x}_k, T^*T\mathbf{x}_k\rangle + \|T^*T\mathbf{x}_k\|^2 \\
&= \|\mathbf{x}_k\|^2 - 2\operatorname{Re}\langle T\mathbf{x}_k, T\mathbf{x}_k\rangle + \|T^*T\mathbf{x}_k\|^2 \\
&= \|\mathbf{x}_k\|^2 - 2\|T\mathbf{x}_k\|^2 + \|T^*T\mathbf{x}_k\|^2. \qquad (10.2.4)
\end{aligned}$$

Now use the fact that T is a contraction to obtain

$$\begin{aligned}
\|T^*T\mathbf{x}_k\| &\leqslant \|T^*\|\,\|T\mathbf{x}_k\| \\
&\leqslant \|T^*\|\,\|T\|\,\|\mathbf{x}_k\| \\
&\leqslant \|\mathbf{x}_k\|.
\end{aligned}$$

Apply this inequality to (10.2.4) to see that

$$\|(I - T^*T)\mathbf{x}_k\|^2 \leqslant 2\|\mathbf{x}_k\|^2 - 2\|T\mathbf{x}_k\|^2.$$

If we assume that $\|\mathbf{x}_k\| \to 1$ and $\|T\mathbf{x}_k\| \to 1$, we obtain $\|(I - T^*T)\mathbf{x}_k\| \to 0$.

Conversely, if $\|\mathbf{x}_k\| \to 1$ and $\|(I - T^*T)\mathbf{x}_k\| \to 0$, then

$$|\langle (I - T^*T)\mathbf{x}_k, \mathbf{x}_k\rangle| \leqslant \|(I - T^*T)\mathbf{x}_k\|\,\|\mathbf{x}_k\| \to 0.$$

Hence

$$\|T\mathbf{x}_k\|^2 = \|\mathbf{x}_k\|^2 - \big\langle (I - T^*T)\mathbf{x}_k, \mathbf{x}_k \big\rangle \to 1$$

as required. □

Since we will be computing the operator norm of $B(T)$, we can dispense with the unimodular constant ξ in (10.2.1) and that assume our finite Blaschke product B takes the form

$$B(z) = \prod_{j=1}^{n} \frac{z_j - z}{1 - \overline{z_j} z}.$$

Let

$$b_j(z) = \frac{z_j - z}{1 - \overline{z_j} z}$$

be the jth factor. The following is the key step to proving Theorem 10.2.2.

Lemma 10.2.5 *If T is a contraction, then*

$$\|T b_2(T) \cdots b_n(T)\| = 1 \qquad \Longleftrightarrow \qquad \|b_1(T) b_2(T) \cdots b_n(T)\| = 1.$$

Proof Let $R = b_2(T) \cdots b_n(T)$ and assume that $\|TR\| = 1$. By the definition of the operator norm, there is a sequence $\mathbf{x}_k$ of unit vectors such that $\|TR\mathbf{x}_k\| \to 1$.

Define

$$\mathbf{y}_k = (I - \overline{z_1} T)\mathbf{x}_k$$

and let us first argue that

$$\|b_1(T) R \mathbf{y}_k\|^2 - \|\mathbf{y}_k\|^2 \to 0. \tag{10.2.6}$$

Indeed,

$$\begin{aligned}
\|b_1(T) R \mathbf{y}_k\|^2 - \|\mathbf{y}_k\|^2 &= \|(T - z_1 I) R \mathbf{x}_k\|^2 - \|(I - \overline{z_1} T)\mathbf{x}_k\|^2 \\
&= \|TR\mathbf{x}_k\|^2 - 2\operatorname{Re}(\overline{z_1} \langle TR\mathbf{x}_k, R\mathbf{x}_k\rangle) + |z_1|^2 \|R\mathbf{x}_k\|^2 \\
&\quad - \|\mathbf{x}_k\|^2 + 2\operatorname{Re}(z_1 \langle \mathbf{x}_k, T\mathbf{x}_k\rangle) - |z_1|^2 \|T\mathbf{x}_k\|^2 \\
&= (\|TR\mathbf{x}_k\|^2 - \|\mathbf{x}_k\|^2) - 2\operatorname{Re}\left[z_1(\langle R\mathbf{x}_k, TR\mathbf{x}_k\rangle - \langle \mathbf{x}_k, T\mathbf{x}_k\rangle)\right] \\
&\quad + |z_1|^2 (\|R\mathbf{x}_k\|^2 - \|T\mathbf{x}_k\|^2) \\
&= (\|TR\mathbf{x}_k\|^2 - \|\mathbf{x}_k\|^2) - 2\operatorname{Re}[z_1 \langle \mathbf{x}_k, (I - R^* R) T\mathbf{x}_k\rangle] \\
&\quad + |z_1|^2 (\|R\mathbf{x}_k\|^2 - \|T\mathbf{x}_k\|^2).
\end{aligned} \tag{10.2.7}$$

Since $\|\mathbf{x}_k\| = 1$ and $\|TR\mathbf{x}_k\| \to 1$ (along with the facts that $\|T\| \leqslant 1$ and $\|R\| \leqslant 1$), we can use the inequalities

$$\|TR\mathbf{x}_k\| \leqslant \|R\mathbf{x}_k\| \leqslant 1$$

and

$$\|TR\mathbf{x}_k\| = \|RT\mathbf{x}_k\| \leqslant \|T\mathbf{x}_k\| \leqslant 1,$$

to conclude that

$$\|R\mathbf{x}_k\| \to 1 \quad \text{and} \quad \|T\mathbf{x}_k\| \to 1.$$

Combine this with the fact that $\|RT\mathbf{x}_k\| \to 1$, along with Lemma 10.2.3, to see that $\|(I - R^*R)T\mathbf{x}_k\| \to 0$. Substitute this limit along with the limits

$$\|\mathbf{x}_k\| \to 1, \quad \|TR\mathbf{x}_k\| \to 1, \quad \|R\mathbf{x}_k\| \to 1, \quad \|T\mathbf{x}_k\| \to 1$$

into (10.2.7) to conclude that $\|b_1(T)R\mathbf{y}_k\|^2 - \|\mathbf{y}_k\|^2 \to 0$, which proves (10.2.6).

Use the inequality

$$\begin{aligned} 1 &= \|\mathbf{x}_k\| \\ &= \|(I - \overline{z_1}T)^{-1}(I - \overline{z_1}T)\mathbf{x}_k\| \\ &\leqslant \|(I - \overline{z_1}T)^{-1}\| \, \|(I - \overline{z_1}T)\mathbf{x}_\mathbf{k}\| \\ &= \|(I - \overline{z_1}T)^{-1}\| \, \|\mathbf{y}_k\| \end{aligned}$$

to obtain

$$\|\mathbf{y}_k\| \geqslant \frac{1}{\|(I - \overline{z_1}T)^{-1}\|} > 0.$$

Using (10.2.6) we get

$$\left\| b_1(T)R\frac{\mathbf{y}_k}{\|\mathbf{y}_k\|} \right\| - 1 \to 0$$

and hence $\|b_1(T)R\| = 1$.

Conversely, assume that

$$\|b_1(T)b_2(T) \cdots b_n(T)\| = 1.$$

Let

$$\psi(z) = \frac{z + z_1}{1 + \overline{z_1} z},$$

which equals the inverse of b_1, and define

$$\psi_j(z) = b_j \circ \psi, \quad 2 \leqslant j \leqslant n.$$

If $T_1 = b_1(T)$, then $b_j(T) = \psi_j(T_1)$ for $2 \leqslant j \leqslant n$. Since

$$\|b_1(T)b_2(T) \cdots b_n(T)\| = 1 \iff \|T_1\psi_2(T_1) \cdots \psi_n(T_1)\| = 1,$$

the previous argument shows that

$$\|\psi_1(T_1) \cdots \psi_n(T_1)\| = 1 \implies \|T_1\psi_2(T_1) \cdots \psi_n(T_1)\| = 1.$$

The proof is now complete. □

To prove Theorem 10.2.2, apply Lemma 10.2.5 n times.

10.3 Numerical Range

Let $\sigma(T)$ denote the spectrum of $T \in \mathscr{L}(\mathscr{H})$. For each $T \in \mathscr{L}(\mathscr{H})$ and analytic polynomial p, the operator $p(T)$ is well defined. The Spectral Mapping Theorem (Theorem A.7.6) asserts that $\sigma(p(T)) = p(\sigma(T))$. See Appendix A.6 for a brief review of operator spectra.

Although there is no spectral mapping theorem for the numerical range (defined below), there are some substitutes from work of Halmos, Berger, Stampfli, and Drury in which finite Blaschke products come into play. We first require a few facts about the numerical range.

Definition 10.3.1 For $T \in \mathscr{L}(\mathscr{H})$,

$$W(T) := \{\langle T\mathbf{x}, \mathbf{x}\rangle : \mathbf{x} \in \mathscr{H}, \|\mathbf{x}\| = 1\}$$

is the *numerical range* of T.

Proposition 10.3.2 *Let $T \in \mathscr{L}(\mathscr{H})$.*

(a) $W(T) \subseteq \{z : |z| \leqslant \|T\|\}$.
(b) If $\mathscr{H}$ is finite dimensional, then $W(T)$ is compact.
(c) If U is unitary, then $W(U^*TU) = W(T)$.
(d) If $S = aT + bI$, in which $a, b \in \mathbb{C}$, then

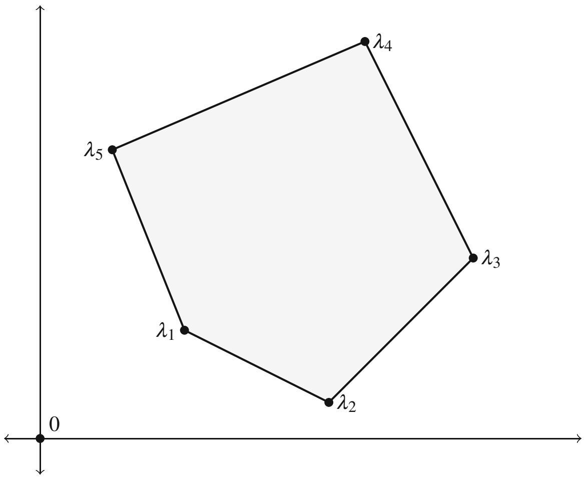


Fig. 10.1 The numerical range of a normal matrix is the convex hull of its eigenvalues λ_j

$$W(S) = aW(T) + b. \tag{10.3.3}$$

(e) $W(T)$ *contains the eigenvalues of* T.
(f) *If* $T \in M_n$ *is normal, then* $W(T)$ *is the convex hull of its eigenvalues.*

The proof of the preceding is left to the reader; see Exercise 10.8. Proposition 10.3.2.f is illustrated in Fig. 10.1.

Definition 10.3.4 The *numerical radius* of $T \in \mathscr{L}(\mathscr{H})$ is

$$w(T) := \sup\{|\langle T\mathbf{x}, \mathbf{x}\rangle| : \mathbf{x} \in \mathscr{H},\ \|\mathbf{x}\| = 1\}.$$

By (10.3.3) we see that

$$w(\lambda T) = |\lambda| w(T), \qquad \lambda \in \mathbb{C}, \tag{10.3.5}$$

and

$$w(S + T) \leqslant w(S) + w(T), \qquad S, T \in \mathscr{L}(\mathscr{H}). \tag{10.3.6}$$

The numerical radius is related to the operator norm via the following inequalities.

Lemma 10.3.7 *If* $T \in \mathscr{L}(\mathscr{H})$, *then*

$$\frac{\|T\|}{2} \leqslant w(T) \leqslant \|T\|.$$

Proof For any unit vector $\mathbf{x} \in \mathscr{H}$, the Cauchy–Schwarz inequality and the definition of the operator norm, yield

$$w(T) = \sup_{\|\mathbf{x}\|=1} |\langle T\mathbf{x}, \mathbf{x}\rangle| \leqslant \sup_{\|\mathbf{x}\|=1} \|T\mathbf{x}\|\|\mathbf{x}\| \leqslant \|T\|.$$

For the lower inequality, we need the polarization identity

$$\begin{aligned} 4\langle T\mathbf{x}, \mathbf{y}\rangle &= \langle T(\mathbf{x}+\mathbf{y}), (\mathbf{x}+\mathbf{y})\rangle - \langle T(\mathbf{x}-\mathbf{y}), (\mathbf{x}-\mathbf{y})\rangle \\ &\quad + i\langle T(\mathbf{x}+i\mathbf{y}), (\mathbf{x}+i\mathbf{y})\rangle - i\langle T(\mathbf{x}-i\mathbf{y}), (\mathbf{x}-i\mathbf{y})\rangle. \end{aligned}$$

This implies

$$\begin{aligned} 4|\langle T\mathbf{x}, \mathbf{y}\rangle| &\leqslant |\langle T(\mathbf{x}+\mathbf{y}), (\mathbf{x}+\mathbf{y})\rangle| + |\langle T(\mathbf{x}-\mathbf{y}), (\mathbf{x}-\mathbf{y})\rangle| \\ &\quad + |\langle T(\mathbf{x}+i\mathbf{y}), (\mathbf{x}+i\mathbf{y})\rangle| + |\langle T(\mathbf{x}-i\mathbf{y}), (\mathbf{x}-i\mathbf{y})\rangle|. \end{aligned}$$

From the definition of $w(T)$ we get the inequality

$$|\langle T\mathbf{z}, \mathbf{z}\rangle| \leqslant w(T)\|\mathbf{z}\|^2$$

for any $\mathbf{z} \in \mathscr{H}$. Apply this estimate to the previous line to get

$$\begin{aligned} 4|\langle T\mathbf{x}, \mathbf{y}\rangle| &\leqslant w(T)\|\mathbf{x}+\mathbf{y}\|^2 + w(T)\|\mathbf{x}-\mathbf{y}\|^2 \\ &\quad + w(T)\|\mathbf{x}+i\mathbf{y}\|^2 + w(T)\|\mathbf{x}-i\mathbf{y}\|^2. \end{aligned}$$

For unit vectors $\mathbf{x}, \mathbf{y}$, two applications of the parallelogram identity

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$$

provide

$$\begin{aligned} 4|\langle T\mathbf{x}, \mathbf{y}\rangle| &\leqslant w(T)(\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 + \|\mathbf{x}+i\mathbf{y}\|^2 + \|\mathbf{x}-i\mathbf{y}\|^2) \\ &= 4w(T)(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) \\ &\leqslant 8w(T). \end{aligned}$$

Thus,

$$\|T\| = \sup\{|\langle T\mathbf{x}, \mathbf{y}\rangle| : \|\mathbf{x}\| \leqslant 1, \|\mathbf{y}\| \leqslant 1\} \leqslant 2w(T). \quad \square$$

Both of the inequalities in Lemma 10.3.7 can be attained; see Exercise 10.12.

Corollary 10.3.8 *If $S, T \in \mathscr{L}(\mathscr{H})$, then*

$$|w(S) - w(T)| \leqslant \|S - T\|.$$

Proof The subadditivity of the numerical radius from (10.3.6) shows that

$$|w(S) - w(T)| \leqslant w(S - T).$$

Now apply Lemma 10.3.7. □

Since the expression $\langle T\mathbf{x}, \mathbf{y}\rangle$ involves only two vectors, it is often fruitful to consider the compression of T onto the two-dimensional subspace spanned by $\mathbf{x}$ and $\mathbf{y}$. This reduces a potentially infinite-dimensional problem to a two-dimensional problem. Our proof of the following seminal result of Hausdorff [69] and Toeplitz [133] employs this strategy. After the initial reduction, our proof largely follows the well-known matrix case [96].

Theorem 10.3.9 (Hausdorff–Toeplitz) *$W(T)$ is convex for all $T \in \mathscr{L}(\mathscr{H})$.*

Proof Let $a, b \in W(T)$. Then there are unit vectors $\mathbf{x}, \mathbf{y} \in \mathscr{H}$ such that

$$\langle T\mathbf{x}, \mathbf{x}\rangle = a \quad \text{and} \quad \langle T\mathbf{y}, \mathbf{y}\rangle = b.$$

We wish to show that the line segment $[a, b]$ connecting the complex numbers a and b remains in $W(T)$. If $\mathbf{x}$ and $\mathbf{y}$ are scalar multiples of each other, then the result is immediate. Hence we assume that $\mathbf{x}$ and $\mathbf{y}$ are linearly independent. Let

$$\mathscr{K} = \operatorname{span}\{\mathbf{x}, \mathbf{y}\}$$

and let P denote the orthogonal projection of $\mathscr{H}$ onto $\mathscr{K}$. Then

$$a = \langle T\mathbf{x}, \mathbf{x}\rangle = \langle PTP\mathbf{x}, \mathbf{x}\rangle \qquad \text{and} \qquad b = \langle T\mathbf{y}, \mathbf{y}\rangle = \langle PTP\mathbf{y}, \mathbf{y}\rangle.$$

Consider the compression

$$T' = PTP|_{\mathscr{K}}$$

of T to $\mathscr{K}$ observe that $W(T') \subseteq W(T)$. In fact, for each $\mathbf{z} \in \mathscr{K}$,

$$\begin{aligned}\langle T'\mathbf{z}, \mathbf{z}\rangle_{\mathscr{K}} &= \langle PTP\mathbf{z}, \mathbf{z}\rangle_{\mathscr{H}}\\ &= \langle TP\mathbf{z}, P\mathbf{z}\rangle_{\mathscr{H}}\\ &= \langle T\mathbf{z}, \mathbf{z}\rangle_{\mathscr{H}}.\end{aligned}$$

Thus, it suffices to show that $[a, b] \subseteq W(T')$. Because T' is a linear transformation from a two-dimensional space to itself, it suffices to prove that the numerical range of a 2×2 matrix is convex.

Suppose that $T \in M_2$. If T is normal, then Proposition 10.3.2 implies that $W(T)$ is convex. If T is not normal, then the matrix

$$T - \tfrac{1}{2}\,\mathrm{tr}(T)I$$

has trace zero and, by (10.3.3), its numerical range is convex if and only if the numerical range of T is convex. Thus, we may assume that T has trace zero. Since the trace is invariant under unitary equivalence, Schur's theorem on unitary triangularization (Theorem A.8.1) implies that T is unitarily equivalent to

$$\begin{bmatrix} \alpha & \beta \\ 0 & -\alpha \end{bmatrix} \tag{10.3.10}$$

for some $\alpha, \beta \in \mathbb{C}$. Moreover, $\beta \neq 0$ since we are assuming that T is not normal.

If $\alpha = 0$, then a computation with the Arithmetic-Geometric mean inequality (see (10.4.3) below) confirms that $W(T)$ is a disk about the origin of radius $|\beta|/2$. If $\alpha \neq 0$, we may appeal to (10.3.3) and replace the matrix in (10.3.10) by

$$\begin{bmatrix} 1 & 2\gamma \\ 0 & -1 \end{bmatrix},$$

in which $\gamma > 0$. Another computation shows that the numerical range of this matrix is an ellipse with principal axes of lengths 2γ and $2\sqrt{1+\gamma^2}$. In both cases, $W(T)$ is convex. □

Corollary 10.3.11 *$\sigma(T) \subseteq W(T)^-$ for all $T \in \mathscr{L}(\mathscr{H})$.*

Proof If $\lambda \in \partial\sigma(T)$, then λ belongs to the approximate point spectrum of T [27, VII.6.7] and hence there are unit vectors $\mathbf{x}_n \in \mathscr{H}$ such that $\langle T\mathbf{x}_n, \mathbf{x}_n\rangle \to \lambda$ as $n \to \infty$. Therefore, $\lambda \in W(T)^-$. If $\lambda \in \sigma(T)\backslash\partial\sigma(T)$, then any line through λ intersects $\partial\sigma(T)$ in at least two points, say λ_1, λ_2. Indeed, let

$$t_+ = \sup\{t \in \mathbb{R} : t\lambda \in \sigma(T)\} \quad \text{and} \quad t_- = \inf\{t \in \mathbb{R} : t\lambda \in \sigma(T)\},$$

then verify that $t_-\lambda$ and $t_+\lambda$ belong to $\partial\sigma(T)$. The preceding discussion ensures that $\lambda_1, \lambda_2 \in W(T)^-$. Since $W(T)$ is convex (Theorem 10.3.9) and since the closure of a convex set is also convex (see Exercise 10.10), it follows that $W(T)^-$ is convex. Thus, $\lambda \in W(T)^-$. □

10.4 Halmos' Conjecture

Is there a version of the spectral mapping theorem for the numerical range? In other words, is

$$W(p(T)) = p(W(T))$$

for analytic polynomials p and $T \in \mathcal{L}(\mathcal{H})$? The following simple example shows that in this generality the conjecture is not true.

Example 10.4.1 We claim that $W(T) = \mathbb{D}^-$ when

$$T = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}. \tag{10.4.2}$$

Indeed, for all unit vectors $\mathbf{x} = (x_1, x_2) \in \mathbb{C}^2$, the Arithmetic-Geometric Mean inequality implies that

$$|\langle T\mathbf{x}, \mathbf{x}\rangle| = 2|\overline{x_1}x_2| \leqslant |x_1|^2 + |x_2|^2 \leqslant 1. \tag{10.4.3}$$

Thus, $W(T) \subseteq \mathbb{D}^-$. If $x_1 = e^{-i\alpha}\cos\theta$, $x_2 = \sin\theta$, and $\alpha, \theta \in [0, 2\pi]$, then

$$\langle T\mathbf{x}, \mathbf{x}\rangle = e^{i\alpha}\sin 2\theta.$$

Let α and θ vary in $[0, 2\pi]$ and conclude that $W(T) = \mathbb{D}^-$. However, since $T^2 = 0$, it follows that $W(T^2) = \{0\}$ and hence the naïve conjecture

$$W(p(T)) = p(W(T))$$

fails for the polynomial $p(z) = z^2$.

Despite this shortcoming, some weaker results hold. Halmos conjectured that

$$T \in \mathcal{L}(\mathcal{H}) \quad \text{and} \quad W(T) \subseteq \mathbb{D}^- \quad \implies \quad W(T^n) \subseteq \mathbb{D}^-, \quad n \geqslant 1.$$

By (10.3.5) and Lemma 10.3.7, this conjecture translates into

$$w(T^n) \leqslant w(T)^n, \quad n \geqslant 1. \tag{10.4.4}$$

We prove a more general version of the conjecture in which the map $z \mapsto z^n$ is replaced with certain finite Blaschke products. Our proof follows [88].

Theorem 10.4.5 *Let $T \in \mathcal{L}(\mathcal{H})$, $w(T) \leqslant 1$, and let B be a finite Blaschke product with $B(0) = 0$. Then $w(B(T)) \leqslant 1$.*

Proof Since $w(T) \leqslant 1$, we have $W(T) \subseteq \mathbb{D}^-$ and hence $\sigma(T) \subseteq \mathbb{D}^-$ by Corollary 10.3.11. We first suppose that $\sigma(T) \subseteq \mathbb{D}$; the general case will be handled by an approximation argument. Theorem 10.1.16 ensures that $B(T)$ is well defined and the spectral mapping theorem (Theorem A.7.6) implies that

$$\sigma(B(T)) = B(\sigma(T)) \subseteq \mathbb{D}.$$

Let $\mathbf{x} \in \mathscr{H}$ with $\|\mathbf{x}\| = 1$. Since $B(0) = 0$, for each $\alpha \in \mathbb{T}$, Corollary 5.2.10 provides $\zeta_1, \zeta_2 \ldots, \zeta_n \in \mathbb{T}$ and $c_1, c_2, \ldots, c_n > 0$ so that

$$\frac{1}{1 - \overline{\alpha} B(z)} = \frac{c_1}{1 - \overline{\zeta}_1 z} + \cdots + \frac{c_n}{1 - \overline{\zeta}_n z}. \tag{10.4.6}$$

In particular, each $I - \overline{\zeta_k} T$ is invertible because $\sigma(T)$ is contained in $\mathbb{D}$. Then

$$\begin{aligned}
1 - \overline{\alpha}\langle B(T)\mathbf{x}, \mathbf{x}\rangle &= \langle (I - \overline{\alpha} B(T))\mathbf{x}, \mathbf{x}\rangle \\
&= \langle \mathbf{y}, (I - \overline{\alpha} B(T))^{-1}\mathbf{y}\rangle && (\mathbf{y} = (I - \overline{\alpha} B(T))\mathbf{x}) \\
&= \Big\langle \mathbf{y}, \sum_{k=1}^{n} c_k (I - \overline{\zeta}_k T)^{-1}\mathbf{y}\Big\rangle && (\text{by } (10.4.6)) \\
&= \sum_{k=1}^{n} c_k \langle \mathbf{y}, (I - \overline{\zeta}_k T)^{-1}\mathbf{y}\rangle && (c_1, c_2, \ldots, c_n > 0) \\
&= \sum_{k=1}^{n} c_k \langle (I - \overline{\zeta}_k T)\mathbf{z}_k, \mathbf{z}_k\rangle && (\mathbf{z}_k = (I - \overline{\zeta}_k T)^{-1}\mathbf{y}) \\
&= \sum_{k=1}^{n} c_k (\|\mathbf{z}_k\|^2 - \overline{\zeta}_k \langle T\mathbf{z}_k, \mathbf{z}_k\rangle).
\end{aligned}$$

Since $w(T) \leqslant 1$, we have

$$\operatorname{Re}\big(\|\mathbf{z}_k\|^2 - \overline{\zeta}_k \langle T\mathbf{z}_k, \mathbf{z}_k\rangle\big) \geqslant 0,$$

and since $c_k > 0$ for all k, it follows that

$$\operatorname{Re}(1 - \overline{\alpha}\langle B(T)\mathbf{x}, \mathbf{x}\rangle) \geqslant 0.$$

Because the preceding holds for all $\alpha \in \mathbb{T}$ and all unit vectors $\mathbf{x}$, we conclude that $w(B(T)) \leqslant 1$.

Now we relax the assumption that $\sigma(T) \subseteq \mathbb{D}$. Let us first show that

$$\lim_{r \to 1^-} B(rT) = B(T) \tag{10.4.7}$$

in the operator norm. By Corollary 10.3.11, we know that $\sigma(T) \subseteq \mathbb{D}^-$, and hence the spectral radius formula (Theorem A.6.11) implies that

$$\lim_{n \to \infty} \|T^n\|^{1/n} \leqslant 1.$$

Since $B(z) = \sum_{n=0}^{\infty} a_n z^n \in \operatorname{Hol}(\mathbb{D}^-)$, the Cauchy–Hadamard formula for the radius of convergence of a power series implies that

$$\limsup_{n \to \infty} |a_n|^{1/n} < 1.$$

Thus,

$$\|B(rT) - B(T)\| \leqslant \sum_{n=0}^{\infty} |a_n|(1 - r^n)\|T^n\|.$$

The Dominated Convergence Theorem now yields (10.4.7). Since $w(B(rT)) \leqslant 1$ for all $r < 1$, we let $r \to 1^-$ and use to Corollary 10.3.8 to complete the proof. □

10.5 The Wiener Algebra Versus the Disk Algebra

We would like to expand the above discussion beyond the Wiener algebra $\mathscr{W}(\mathbb{D})$ to the disk algebra $\mathscr{A}(\mathbb{D})$ (the set of continuous functions on $\mathbb{D}^-$ that are analytic on $\mathbb{D}$). We have already discussed the definition of $f(T)$ when $f \in \mathscr{W}(\mathbb{D})$. We need to do the same when f belongs to $\mathscr{A}(\mathbb{D})$. Moreover, since the Taylor coefficients of an $f \in \mathscr{W}(\mathbb{D})$ are absolutely summable, the Taylor polynomials of f converge uniformly on $\mathbb{D}^-$ and hence $f \in \mathscr{A}(\mathbb{D})$. It turns out that the containment $\mathscr{W}(\mathbb{D}) \subseteq \mathscr{A}(\mathbb{D})$ is proper.

Theorem 10.5.1 $\mathscr{W}(\mathbb{D}) \subsetneq \mathscr{A}(\mathbb{D})$

The remainder of this section is devoted to the proof of Theorem 10.5.1 which was inspired by du Bois Reymond (see the end notes of this chapter). In order to prove this, we need a sequence of analytic polynomials

$$p_n(z) = \sum_{k=0}^{d_n} a_{n,k} z^k \tag{10.5.2}$$

with the following two properties.

(a) They are uniformly bounded on $\mathbb{D}^-$:

$$|p_n(z)| \leqslant 1, \quad n \geqslant 1, \; z \in \mathbb{D}^-.$$

(b) Their ℓ^1 norms are not uniformly bounded:

$$\lim_{n \to \infty} \sum_{k=0}^{d_n} |a_{n,k}| = \infty.$$

Toward the end of this section, we present an explicit construction. Assuming the existence of such a sequence of polynomials, we construct f as follows. According to (a), we can replace the sequence p_n by a subsequence so that

$$\sum_{n=1}^{\infty} \frac{1}{n^2}\Big(\sum_{k=0}^{d_n} |a_{n,k}|\Big) = \infty. \tag{10.5.3}$$

Then we multiply each polynomial p_n by an appropriate monomial z^{k_n} so that the nonvanishing coefficients of the new sequence never coincide. For example, we can take $k_1 = 1$ and

$$k_n = k_1 + \cdots + k_{n-1} + n. \tag{10.5.4}$$

Then let

$$f(z) = \sum_{n=1}^{\infty} \frac{1}{n^2} z^{k_n} p_n(z). \tag{10.5.5}$$

Property (a) ensures that $f \in \mathscr{A}(\mathbb{D})$. Property (b), or more precisely the choice (10.5.3), says that $f \notin \mathscr{W}(\mathbb{D})$.

To construct the sequence p_n, which is the heart of the construction above, we need a technical lemma (Fig. 10.2).

Lemma 10.5.6 *Let*

$$S_n(\theta) = \sum_{k=1}^{n} \frac{\sin(k\theta)}{k}, \quad n \geqslant 1,\ \theta \in \mathbb{R}.$$

Then $|S_n(\theta)| \leqslant 5$ *for all* $n \geqslant 1$ *and* $\theta \in \mathbb{R}$.

Proof Since S_n is a 2π-periodic odd function, we may assume that $\theta \in (0, \pi)$. If $\lfloor \pi/\theta \rfloor < n$, we write

$$S_n(\theta) = \sum_{k=1}^{\lfloor \pi/\theta \rfloor} \frac{\sin(k\theta)}{k} + \sum_{k=\lfloor \pi/\theta \rfloor+1}^{n} \frac{\sin(k\theta)}{k} = S_n^{(1)}(\theta) + S_n^{(2)}(\theta).$$

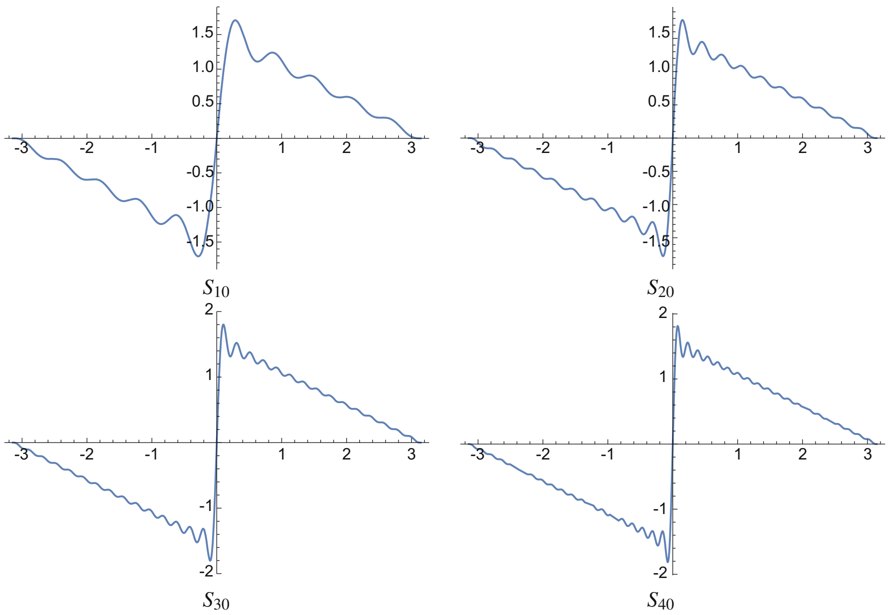


Fig. 10.2 The graph of S_n for $n = 10, 20, 30, 40$

Then

$$\begin{aligned}|S_n^{(1)}(\theta)| &\leqslant \sum_{k=1}^{\lfloor \pi/\theta \rfloor} \frac{|\sin(k\theta)|}{k} \\ &\leqslant \sum_{k=1}^{\lfloor \pi/\theta \rfloor} \frac{k\theta}{k} \\ &= \theta \lfloor \pi/\theta \rfloor \\ &\leqslant \pi\end{aligned}$$

and

$$\begin{aligned}|S_n^{(2)}(\theta)| &= \left| \sum_{k=\lfloor \pi/\theta \rfloor + 1}^{n} \frac{\sin(k\theta)}{k} \right| \\ &\leqslant \frac{1}{\lfloor \pi/\theta \rfloor + 1} \cdot \frac{1}{\sin(\theta/2)} \\ &\leqslant \frac{\pi}{(\lfloor \pi/\theta \rfloor + 1)\theta} \\ &\leqslant 1.\end{aligned}$$

Hence, $|S_n(\theta)| \leqslant 1 + \pi \leqslant 5$: see Exercise 10.15.

If $\lfloor \pi/\theta \rfloor \geqslant n$, then the estimate above for $S_n^{(1)}$ shows that $|S_n(\theta)| \leqslant \pi \leqslant 5$. □

Write S_n as

$$S_n(\theta) = \sum_{k=1}^{n} \left(\frac{i}{2k} e^{-ik\theta} - \frac{i}{2k} e^{ik\theta} \right)$$

to see that S_n is a trigonometric polynomial of degree n. Now define

$$p_n(e^{i\theta}) = \frac{1}{5} e^{in\theta} S_n(\theta). \tag{10.5.7}$$

Then p_n is an analytic polynomial of the form in (10.5.2) with $d_n = 2n$. By Lemma 10.5.6 and the Maximum Modulus Principle, property (a) is fulfilled. Furthermore,

$$\sum_{k=0}^{d_n} |a_{n,k}| = \frac{1}{5} \sum_{k=1}^{n} \frac{1}{k} \asymp \log n \to \infty.$$

Hence, property (b) is also satisfied.

We can get a bit more out of the above construction. Let us, instead of (10.5.3), choose the subsequence such that it satisfies the more restrictive condition

$$\sum_{k=0}^{d_n} |a_{n,k}| \geqslant n^3.$$

For the specific choice of polynomials above, this is equivalent to

$$\log d_n \geqslant n^3. \tag{10.5.8}$$

As we saw before, we have $f \in \mathscr{A}(\mathbb{D}) \backslash \mathscr{W}(\mathbb{D})$. Moreover, the Taylor series of f does not converge uniformly to f on $\mathbb{D}^-$. Even more dramatically, the Taylor series for f diverges at $z = 1$. To see this, write

$$f(z) = \sum_{n=1}^{\infty} \alpha_n z^n$$

and

$$T_N(z) = \sum_{n=1}^{N} \alpha_n z^n.$$

Recall from (10.5.4) that the Taylor coefficients of the polynomials $z^{k_n} p_n(z)$ have disjoint support and $p_n(1) = 0$. Choose N to be the index of the midpoint for Taylor coefficients of the mth polynomial. Then, by (10.5.5) and (10.5.7), we have

$$\begin{aligned} T_N(1) &= \sum_{n=1}^{N} \alpha_n \\ &= \sum_{n=1}^{m-1} \frac{1}{n^2} p_n(1) + \frac{i}{10m^2} \left(\frac{1}{d_m} + \frac{1}{d_m - 1} + \cdots + \frac{1}{2} + 1 \right) \\ &\asymp \frac{\log d_m}{m^2}. \end{aligned}$$

The assumption (10.5.8) now ensures that $T_N(1) = O(m) \to \infty$. Thus, the Taylor series for f diverges at $z = 1$.

10.6 The Berger–Stampfli Mapping Theorem

In this section we present a proof of a theorem of Berger and Stampfli [7] that serves as a generalization of the Halmos conjecture. Since this theorem concerns the numerical range of $f(T)$ for certain $T \in \mathscr{L}(\mathscr{H})$ and f in the disk algebra $\mathscr{A}(\mathbb{D})$, we first need to define $f(T)$. Recall from (10.1.15) that we defined $f(T)$ for f in the Wiener algebra $\mathscr{W}(\mathbb{D})$. However, $\mathscr{W}(\mathbb{D})$ is a proper subset of $\mathscr{A}(\mathbb{D})$ and hence the definition of $f(T)$ for $f \in \mathscr{A}(\mathbb{D})$ requires more care. The proof of the next lemma, which depends on finite Blaschke products in a crucial way, follows [88].

Lemma 10.6.1 *If $f \in \mathscr{A}(\mathbb{D})$, $T \in \mathscr{L}(\mathscr{H})$, and $w(T) \leqslant 1$, then*

$$\lim_{r \to 1^-} f(rT)$$

exists in the operator norm.

Proof Suppose that $T \in \mathscr{L}(\mathscr{H})$ and $w(T) \leqslant 1$. Corollary 10.3.11 implies that $\sigma(T) \subseteq W(T)^- \subseteq \mathbb{D}^-$. Let us first prove that

$$w(g(T)) \leqslant \|g\|_\infty \tag{10.6.2}$$

whenever g is analytic in a neighborhood of $\mathbb{D}^-$, $g(0) = 0$, and $\sigma(T) \subseteq \mathbb{D}$. By scaling, we may also assume that $\|g\|_\infty = 1$. Since $g \in \operatorname{Hol}(\mathbb{D}^-)$ we can employ the argument used to prove (10.4.7) to see that $g(T)$ is a well-defined element of $\mathscr{L}(\mathscr{H})$. Carathéodory's Theorem (Theorem 4.1.1) provides a sequence of finite Blaschke products B_n that converges uniformly on compact subsets of $\mathbb{D}$ to g. Since

$g(0) = 0$, we can also arrange that $B_n(0) = 0$ for all n; see Exercise 4.4. From Theorem 10.4.5, we know that $w(B_n(T)) \leqslant 1$ for all n. The Spectral Mapping Theorem (Theorem A.7.6) says that

$$\sigma(B_n(T)) = B_n(\sigma(T)) \subseteq \mathbb{D} \quad \text{and} \quad \sigma(g(T)) = g(\sigma(T)) \subseteq \mathbb{D}.$$

From here, we can use the Riesz functional calculus (see (A.7.5)) to see that $B_n(T)$ converges in norm to $g(T)$. It follows from Corollary 10.3.8 that $w(g(T)) \leqslant 1$.

We apply the argument above to

$$g_{r,s}(z) = f(rz) - f(sz), \qquad r, s \in (0, 1),$$

which is analytic in a neighborhood of $\mathbb{D}^-$ and vanishes at the origin, and to the operator $g_{r,s}(T)$. For r, s close enough to 1, we have $\|g_{r,s}\|_\infty < 1$ and hence

$$\sigma(g_{r,s}(T)) = g_{r,s}(\sigma(T)) \subseteq \mathbb{D}$$

by the Spectral Mapping Theorem. By what we have already shown,

$$w(g_{r,s}(T)) \leqslant \|g_{r,s}\|_\infty$$

and hence, by Lemma 10.3.7,

$$\|f(rT) - f(sT)\| = \|g_{rs}(T)\| \leqslant 2w(g_{rs}(T)) \leqslant 2\|g_{rs}\|_\infty, \tag{10.6.3}$$

which tends to zero as $r, s \to 1^-$. It follows that for each sequence $r_n \to 1^-$, the corresponding sequence $f(r_n T)$ is Cauchy in $\mathscr{L}(\mathscr{H})$ and hence convergent. The estimate (10.6.3) ensures that the limit is independent of the choice of sequence r_n. Therefore, $\lim_{r\to 1^-} f(rT)$ exists. □

The preceding theorem permits us to define

$$f(T) := \lim_{r\to 1^-} f(rT)$$

for $f \in \mathscr{A}(\mathbb{D})$ and $T \in \mathscr{L}(\mathscr{H})$ with $w(T) \leqslant 1$. Now that $f(T)$ is defined, we state the main theorem of this section.

Theorem 10.6.4 (Berger–Stampfli) *If $f \in \mathscr{A}(\mathbb{D})$, $f(0) = 0$, $T \in \mathscr{L}(\mathscr{H})$, and $w(T) \leqslant 1$, then $w(f(T)) \leqslant \|f\|_\infty$.*

Proof The hypotheses imply that $\sigma(T) \subseteq W(T)^- \subseteq \mathbb{D}^-$. We have already proved the result when $\sigma(T) \subseteq \mathbb{D}$; this is (10.6.2). We now relax the assumption that $\sigma(T) \subseteq \mathbb{D}$. From the proof of the previous lemma, we know that $w(f(rT)) \leqslant \|f\|_\infty$ for all $r \in (0, 1)$. Since $f(T) = \lim_{r\to 1^-} f(rT)$, Corollary 10.3.8 implies that $w(f(T)) \leqslant \|f\|_\infty$. □

What happens if $f(0) \neq 0$ in Theorem 10.6.4? This will be addressed in the next few sections.

10.7 A Local Inequality

Let $T \in \mathscr{L}(\mathscr{H})$ and $\mathbf{x} \in \mathscr{H}$. The left-hand inequality in Lemma 10.3.7 amounts to saying that $\|T\mathbf{x}\| \leqslant 2$ whenever $w(T) \leqslant 1$ and $\|\mathbf{x}\| \leqslant 1$. The following result from [88] is local refinement of this fact.

Theorem 10.7.1 (Mashreghi–Ransford) *If $w(T) \leqslant 1$ and $\|\mathbf{x}\| \leqslant 1$, then*

$$\|T\mathbf{x}\|^2 \leqslant 2 + 2\sqrt{1 - |\langle T\mathbf{x}, \mathbf{x}\rangle|^2}. \tag{10.7.2}$$

Proof Without loss of generality, we may assume that $\|\mathbf{x}\| = 1$ and $\langle T\mathbf{x}, \mathbf{x}\rangle \geqslant 0$. Observe that

$$A = \frac{1}{2}(T + T^*) \quad \text{and} \quad B = \frac{1}{2i}(T - T^*)$$

are self-adjoint and have numerical radius at most 1 by (10.3.6). Consequently, $\|A\| \leqslant 1$ and $\|B\| \leqslant 1$. The condition $\langle T\mathbf{x}, \mathbf{x}\rangle \geqslant 0$ implies that $\langle A\mathbf{x}, \mathbf{x}\rangle = \langle T\mathbf{x}, \mathbf{x}\rangle$ and $\langle B\mathbf{x}, \mathbf{x}\rangle = 0$. This yields

$$\begin{aligned}
\sqrt{\|T\mathbf{x}\|^2 - |\langle T\mathbf{x}, \mathbf{x}\rangle|^2} &= \|T\mathbf{x} - \langle T\mathbf{x}, \mathbf{x}\rangle\mathbf{x}\| \\
&\leqslant \|A\mathbf{x} - \langle A\mathbf{x}, \mathbf{x}\rangle\mathbf{x}\| + \|B\mathbf{x} - \langle B\mathbf{x}, \mathbf{x}\rangle\mathbf{x}\| \\
&= \sqrt{\|A\mathbf{x}\|^2 - |\langle A\mathbf{x}, \mathbf{x}\rangle|^2} + \sqrt{\|B\mathbf{x}\|^2 - |\langle B\mathbf{x}, \mathbf{x}\rangle|^2} \\
&\leqslant \sqrt{1 - |\langle T\mathbf{x}, \mathbf{x}\rangle|^2} + 1,
\end{aligned}$$

which, after some arithmetic, implies (10.7.2). □

From Theorem 10.7.1 we derive the following operator inequality. This result is needed for the proof of Corollary 10.7.6 below.

Corollary 10.7.3 *If $w(T) \leqslant 1$, then*

$$I + 2t(T + T^*) + (t^2 - \tfrac{1}{4})T^*T \geqslant 0, \qquad t \in [0, \tfrac{1}{2}]. \tag{10.7.4}$$

Proof The inequality (10.7.4) is equivalent to

$$1 + 2t\,\mathrm{Re}\langle T\mathbf{x}, \mathbf{x}\rangle + (t^2 - \tfrac{1}{4})\|T\mathbf{x}\|^2 \geqslant 0, \qquad t \in [0, \tfrac{1}{2}], \quad \|\mathbf{x}\| = 1.$$

To prove this, we consider two cases. If $\|T\mathbf{x}\|^2 \leqslant 2$, then for all $t \in [0, \frac{1}{2}]$,

$$\begin{aligned} 1 + 2t\,\mathrm{Re}\langle T\mathbf{x}, \mathbf{x}\rangle + (t^2 - \tfrac{1}{4})\|T\mathbf{x}\|^2 &\geqslant 1 + 2t\,\mathrm{Re}\langle T\mathbf{x}, \mathbf{x}\rangle + 2(t^2 - \tfrac{1}{4}) \\ &= 2\left|t + \frac{\langle T\mathbf{x}, \mathbf{x}\rangle}{2}\right|^2 + \frac{1 - |\langle T\mathbf{x}, \mathbf{x}\rangle|^2}{2} \\ &\geqslant 0. \end{aligned}$$

If $\|T\mathbf{x}\|^2 > 2$, then write (10.7.2) in the form

$$\|T\mathbf{x}\|^2 - 2 \leqslant 2\sqrt{1 - |\langle T\mathbf{x}, \mathbf{x}\rangle|^2}$$

and square both sides to get

$$4\|T\mathbf{x}\|^2 - \|T\mathbf{x}\|^4 - 4|\langle T\mathbf{x}, \mathbf{x}\rangle|^2 \geqslant 0.$$

For all $t \in [0, \frac{1}{2}]$, we have

$$\begin{aligned} &1+2t\,\mathrm{Re}\langle T\mathbf{x}, \mathbf{x}\rangle + (t^2 - \tfrac{1}{4})\|T\mathbf{x}\|^2 \\ &\quad= \|T\mathbf{x}\|^2\left|t + \frac{\langle T\mathbf{x}, \mathbf{x}\rangle}{\|T\mathbf{x}\|^2}\right|^2 + \frac{4\|T\mathbf{x}\|^2 - \|T\mathbf{x}\|^4 - 4|\langle T\mathbf{x}, \mathbf{x}\rangle|^2}{4\|T\mathbf{x}\|^2} \geqslant 0. \qquad \square \end{aligned}$$

For fixed $T \in \mathscr{L}(\mathscr{H})$, let $Q(T, t, s)$ be the operator defined by

$$Q(T, t, s) = I + t(T + T^*) + sT^*T.$$

Definition 10.7.5 Let S denote the set of all $(t, s) \in [0, \infty) \times \mathbb{R}$ such that whenever $\mathscr{H}$ is a Hilbert space and $T \in \mathscr{L}(\mathscr{H})$ with $w(T) \leqslant 1$ we have $Q(T, t, s) \geqslant 0$.

The following corollary characterizes S; see Fig. 10.3. We will need this in order to extend Theorem 10.6.4 to the case when $f(0) \neq 0$.

Corollary 10.7.6 *The region S from Definition 10.7.5 is characterized by the following inequalities:*

$$\begin{cases} s \geqslant t^2 - \frac{1}{4}, & \text{if } 0 \leqslant t \leqslant \frac{1}{2}, \\ s \geqslant 2t - 1, & \text{if } \frac{1}{2} \leqslant t \leqslant 1, \\ s \geqslant t^2, & \text{if } t \geqslant 1. \end{cases}$$

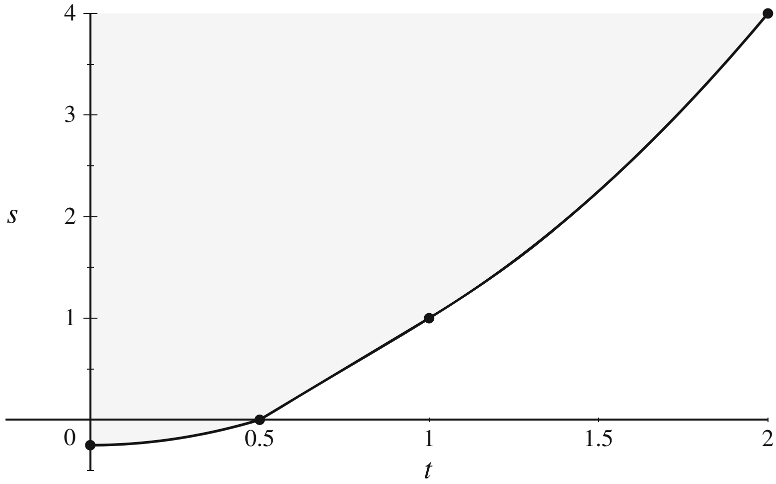


Fig. 10.3 The region S

Proof We divide the argument into three cases, according to the value of t.
Case I ($0 \leqslant t \leqslant \frac{1}{2}$): If $s \geqslant t^2 - \frac{1}{4}$, then Corollary 10.7.3 shows that, for all T with $w(T) \leqslant 1$,

$$Q(T, t, s) \geqslant I + t(T + T^*) + (t^2 - \tfrac{1}{4})T^*T \geqslant 0.$$

On the other hand, if $s < t^2 - \frac{1}{4}$ and

$$T = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix},$$

then $w(T) \leqslant 1$ and

$$Q(T, t, s) = \begin{bmatrix} 1 & 2t \\ 2t & 1 + 4s \end{bmatrix} \not\geqslant 0,$$

since it has negative determinant. Thus, for this range of values of t, we have $(t, s) \in S \iff s \geqslant t^2 - \frac{1}{4}$.
Case II ($\frac{1}{2} \leqslant t \leqslant 1$): If $s \geqslant 2t - 1$, then, for all T with $w(T) \leqslant 1$,

$$\begin{aligned} Q(T, t, s) &\geqslant I + t(T + T^*) + (2t - 1)T^*T \\ &= (1 - t)(2I - (T + T^*)) + (2t - 1)(I + T)^*(I + T) \geqslant 0. \end{aligned}$$

On the other hand, if $s < 2t - 1$ and $T = -I$, then $w(T) \leqslant 1$ and

$$Q(T, t, s) = (1 - 2t + s)I \ngeqslant 0.$$

Therefore, for this range of values of t, we have $(t, s) \in S \iff s \geqslant 2t - 1$.
Case III ($t \geqslant 1$): If $s \geqslant t^2$, then, for all T with $w(T) \leqslant 1$,

$$\begin{aligned} Q(T, t, s) &\geqslant I + t(T + T^*) + t^2 T^* T \\ &= (I + tT)^*(I + tT) \geqslant 0. \end{aligned}$$

On the other hand, if $t \leqslant s < t^2$ and $T = -(t/s)I$, then $w(T) \leqslant 1$ and

$$Q(T, t, s) = (1 - t^2/s)I \ngeqslant 0.$$

Thus, for this range of values of t, we have $(t, s) \in S \iff s \geqslant t^2$. □

10.8 Teardrops and Drury's Theorem

We can formulate the Berger–Stampfli theorem as a numerical range mapping theorem: if $f : \mathbb{D}^- \to \mathbb{D}^-$ belongs to $\mathscr{A}(\mathbb{D})$ and $f(0) = 0$, then

$$W(T) \subseteq \mathbb{D}^- \quad \Longrightarrow \quad W(f(T)) \subseteq \mathbb{D}^-.$$

If $f(0) \neq 0$, the preceding implication may fail; see Sect. 10.9. In this case, the best result is a theorem due to Drury [36]. To state his result, we need to introduce some terminology.

Definition 10.8.1 (Drury's Teardrop Region) For $\alpha \in \mathbb{D}^-$,

$$\operatorname{td}(\alpha) := \operatorname{conv}\left(\mathbb{D}^- \cup D(\alpha, 1 - |\alpha|^2)^-\right)$$

is a *teardrop region*.

The region $\operatorname{td}(\alpha)$ is the convex hull of the union of the closed unit disk and the closed disk of center α and radius $1 - |\alpha|^2$; see Fig. 10.4. When $\alpha \in [0, 1)$, $\operatorname{td}(\alpha)$ also equals the intersection of the two families of half planes

$$\{z : \operatorname{Re}(e^{-i\theta} z) \leqslant 1\}, \qquad \cos\theta \leqslant \alpha, \tag{10.8.2}$$

and

$$\{z : \operatorname{Re}(e^{-i\theta}(z - \alpha)) \leqslant 1 - \alpha^2\}, \qquad \cos\theta \geqslant \alpha; \tag{10.8.3}$$

see Exercise 10.14 and Fig. 10.5. Drury's theorem can now be stated as follows.

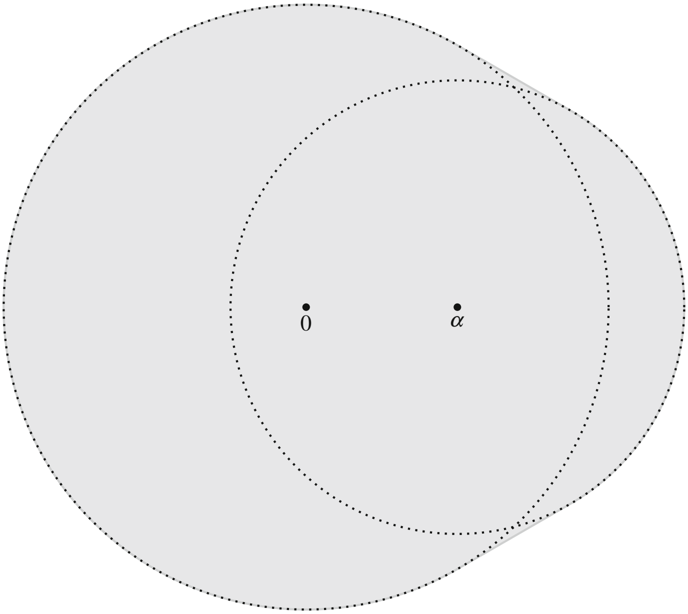


Fig. 10.4 The Drury teardrop region td(α)

Theorem 10.8.4 *Let* $T \in \mathscr{L}(\mathscr{H})$*,* $W(T) \subseteq \mathbb{D}^-$*, and let* $f : \mathbb{D}^- \to \mathbb{D}^-$ *be a function in* $\mathscr{A}(\mathbb{D})$*. Then*

$$W(f(T)) \subseteq \mathrm{td}(f(0)).$$

Proof We follow [36], with a few details added. Let $\alpha = f(0)$. We can assume that $|\alpha| < 1$, since otherwise, by the Maximum Modulus Principle, f is constant and there is nothing to prove. Let ϕ_α be the disk automorphism

$$\phi_\alpha(z) = \frac{\alpha + z}{1 + \overline{\alpha} z}$$

and set $g = \phi_\alpha^{-1} \circ f$. Then g belongs to the disk algebra, $\|g\|_\infty \leqslant 1$, and $g(0) = 0$. By Theorem 10.6.4, we have $W(g(T)) \subseteq \mathbb{D}^-$. Since $f = \phi_\alpha \circ g$, we may proceed by replacing T by $g(T)$ and just study the case $f = \phi_\alpha$. Since $\phi_\alpha(T) = \phi_{|\alpha|}(e^{-i \arg \alpha} T)$, we may also assume that $\alpha \in [0, 1)$.

Because td(α) is the intersection of the two families of half planes (10.8.2) and (10.8.3) (see Exercise 10.14), to show that $W(\phi_\alpha(T)) \subseteq \mathrm{td}(\alpha)$, it suffices to prove that

$$\mathrm{Re}(e^{-i\theta} \phi_\alpha(T)) \leqslant I, \qquad \cos\theta \leqslant \alpha, \tag{10.8.5}$$

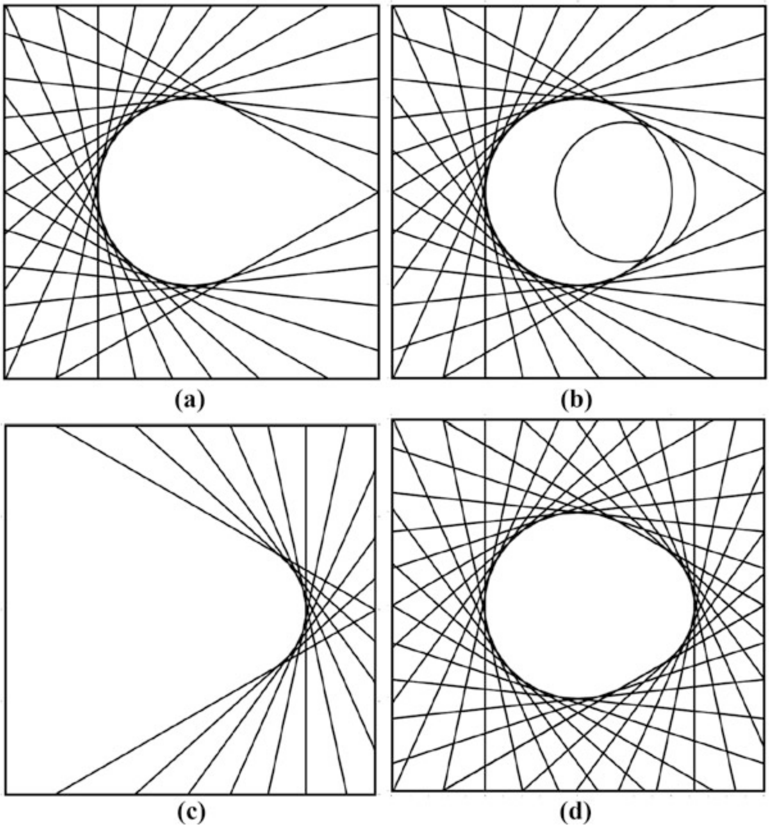


Fig. 10.5 (**a**) is the intersection of the half planes $\{z : \operatorname{Re}(e^{-i\theta}) \leqslant 1\}$ with $\cos\theta \leqslant \alpha$. (**b**) is the intersection of the half planes $\{z : \operatorname{Re}(e^{-i\theta}) \leqslant 1\}$ with $\cos\theta \leqslant \alpha$ together with the circles $|z| = 1$ and $|z - \alpha| \leqslant 1 - |\alpha|^2$. (**c**) is the intersection of the half planes $\{z : \operatorname{Re}(e^{-i\theta}(z - \alpha) \leqslant 1 - |\alpha|^2\}$ with $\cos\theta \geqslant \alpha$. (**d**) is the intersection of the two families of half planes (which form the Drury teardrop region)

and

$$\operatorname{Re}(e^{-i\theta}(\phi_\alpha(T) - \alpha I)) \leqslant (1 - \alpha^2)I, \qquad \cos\theta \geqslant \alpha. \tag{10.8.6}$$

We begin by proving (10.8.5), which is equivalent to

$$2I - e^{-i\theta}\phi_\alpha(T) - e^{i\theta}\phi_\alpha(T^*) \geqslant 0. \tag{10.8.7}$$

If $A, B \in \mathscr{L}(\mathscr{H})$ and B is invertible, then

$$A \geqslant 0 \iff \langle A\mathbf{x}, \mathbf{x}\rangle \geqslant 0 \text{ for all } \mathbf{x} \in \mathscr{H}$$

$$\begin{aligned}
&\iff \langle AB\mathbf{y}, B\mathbf{y}\rangle \geqslant 0 \text{ for all } \mathbf{y} \in \mathscr{H} \\
&\iff \langle B^*AB\mathbf{y}, \mathbf{y}\rangle \geqslant 0 \text{ for all } \mathbf{y} \in \mathscr{H} \\
&\iff B^*AB \geqslant 0.
\end{aligned}$$

Applying this with A equal to the left-hand side of (10.8.7) and $B = (I + \alpha T)$, we see that the desired inequality (10.8.5) is equivalent to

$$\begin{aligned}
&2(1-\alpha\cos\theta)I + (2\alpha - e^{i\theta} - \alpha^2 e^{-i\theta})T \\
&\qquad + (2\alpha - e^{-i\theta} - \alpha^2 e^{i\theta})T^* + 2\alpha(\alpha - \cos\theta)T^*T \geqslant 0.
\end{aligned}$$

If we let

$$\omega = \frac{2\alpha - e^{i\theta} - \alpha^2 e^{-i\theta}}{|2\alpha - e^{i\theta} - \alpha^2 e^{-i\theta}|} = \frac{2\alpha - e^{i\theta} - \alpha^2 e^{-i\theta}}{1 - 2\alpha\cos\theta + \alpha^2},$$

then we may rewrite the last inequality as

$$\begin{aligned}
&2(1-\alpha\cos\theta)I + (1 - 2\alpha\cos\theta + \alpha^2)(\omega T + (\omega T)^*) \\
&\qquad + 2\alpha(\alpha - \cos\theta)(\omega T)^*(\omega T) \geqslant 0,
\end{aligned}$$

or equivalently, as $Q(\omega T, t, s) \geqslant 0$, in which

$$t = \frac{1 - 2\alpha\cos\theta + \alpha^2}{2(1-\alpha\cos\theta)} \qquad \text{and} \qquad s = \frac{\alpha(\alpha - \cos\theta)}{1 - \alpha\cos\theta} = 2t - 1.$$

For $-1 \leqslant \cos\theta \leqslant \alpha$, one can show that $t \in [\frac{1}{2}, 1]$. By Corollary 10.7.6, we have $Q(\omega T, t, s) \geqslant 0$. This establishes (10.8.5).

Now we turn to (10.8.6), which is equivalent to

$$2I - e^{-i\theta}\psi_\alpha(T) - e^{i\theta}\psi_\alpha(T^*) \geqslant 0,$$

where

$$\psi_\alpha(z) = \frac{z}{1 + \alpha z}.$$

As before, when considering B^*AB with $B = (I + \alpha T)$, the preceding inequality is equivalent to

$$2I + (2\alpha - e^{-i\theta})T + (2\alpha - e^{i\theta}T^*) + 2\alpha(\alpha - \cos\theta)T^*T \geqslant 0.$$

If we let

$$\omega = \frac{2\alpha - e^{-i\theta}}{|2\alpha - e^{-i\theta}|} = \frac{2\alpha - e^{-i\theta}}{2\sqrt{\alpha(\alpha - \cos\theta) + \frac{1}{4}}},$$

then we may rewrite the last inequality as

$$I + \sqrt{\alpha(\alpha - \cos\theta) + \tfrac{1}{4}}\big(\omega T + (\omega T)^*\big) + \alpha(\alpha - \cos\theta)(\omega T)^*(\omega T) \geqslant 0,$$

or equivalently, as $Q(\omega T, t, s) \geqslant 0$, in which

$$t = \sqrt{\alpha(\alpha - \cos\theta) + \tfrac{1}{4}} \qquad \text{and} \qquad s = \alpha(\alpha - \cos\theta) = t^2 - \tfrac{1}{4}.$$

For $\alpha \leqslant \cos\theta \leqslant 1$, one can show that $t \in [0, \frac{1}{2}]$. By Corollary 10.7.6, we have $Q(\omega T, t, s) \geqslant 0$. This establishes (10.8.6) and completes the proof. □

The part of the numerical range of $f(T)$ "sticking out" of the unit disk is governed by the inequality (10.8.6), which corresponds to the slice of S (see Definition 10.7.5) where $0 \leqslant t \leqslant \frac{1}{2}$, which is, in turn, determined by the operator inequality in Corollary 10.7.3.

Corollary 10.8.8 *Let $T \in \mathscr{L}(\mathscr{H})$ and $W(T) \subseteq \mathbb{D}^-$, and let $f : \mathbb{D} \to \mathbb{D}$ belong to $\mathscr{A}(\mathbb{D})$. Then*

$$w(f(T)) \leqslant 1 + |f(0)| - |f(0)|^2 \leqslant \tfrac{5}{4}.$$

10.9 Sharpness of Drury's Result via Disk Automorphisms

The optimality of the teardrop region in Theorem 10.8.4 is established by an example. To set the stage, let us properly formulate the question: find the smallest convex set Ω such that

$$f(T) \subseteq \Omega$$

for all functions $f \in \mathscr{A}(\mathbb{D})$ with $f(0) = \alpha$ fixed.

If $\alpha = 0$, thanks to the Halmos conjecture and simple examples (diagonal 2×2 matrices or the example below), we know that $\Omega = \mathbb{D}^-$. Hence, for the rest of discussion, we may assume that $\alpha \neq 0$. The function

$$f(z) = \frac{\alpha + z}{1 + \overline{\alpha} z}$$

belongs to the disk algebra and $f(0) = \alpha$. We consider two classes of operators to show that the two disks in the definition of $\mathrm{td}(\alpha)$ have to be in Ω. Thus, the Drury teardrop region $\mathrm{td}(\alpha)$ is optimal.

Example 10.9.1 Let

$$T = \begin{bmatrix} \zeta & 0 \\ 0 & -\alpha \end{bmatrix},$$

in which $\zeta \in \mathbb{T}$. A computation confirms that $W(T) = [-\alpha, \zeta] \subseteq \mathbb{D}^-$. Moreover,

$$f(T) = \begin{bmatrix} f(\zeta) & 0 \\ 0 & 0 \end{bmatrix}$$

and hence $W(f(T)) = [0, f(\zeta)]$. As ζ runs once through the unit circle $\mathbb{T}$, the argument principle ensures that $f(\zeta)$ does as well. Thus, $\mathbb{D}^- \subseteq \Omega$.

Example 10.9.2 Consider the matrix T given by (10.4.2); that is,

$$T = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}.$$

Example 10.4.1 tells us that $W(T) = \mathbb{D}^-$ and a computation shows

$$f(T) = \begin{bmatrix} \alpha & 2(1 - |\alpha|^2) \\ 0 & \alpha \end{bmatrix}.$$

To find the numerical range of $f(T)$, it is better to write it as

$$f(T) = \alpha I + (1 - |\alpha|^2)T.$$

This identity reveals that

$$\begin{aligned} W(f(T)) &= \alpha + (1 - |\alpha|^2)W(T) \\ &= \alpha + (1 - |\alpha|^2)\mathbb{D}^- \\ &= D(\alpha, 1 - |\alpha|^2)^-. \end{aligned}$$

Therefore, $D(\alpha, 1 - |\alpha|^2)^- \subseteq \Omega$.

10.10 Notes

von Neumann's Inequality in Finite Dimensions

Our presentation of von Neumann's inequality (Theorem 10.1.11) depended on the dilation theorem (Theorem 10.1.5) which is an infinite-dimensional proof. When the contraction T is on a finite-dimensional space (and thus can be considered as a matrix), there are proofs that use either a finite-dimensional version of the dilation theorem [40, 95] or the singular value decomposition [115, Ch. 1].

More on the Numerical Range

Lax and Wendroff [93, 94] showed that if $\mathscr{H}$ is finite dimensional, then $w(T) \leqslant 1$ implies that for some $M > 0$, we have $\|T^n\| \leqslant M$ for all $n \geqslant 1$. However, their method is such that the upper bound M depends on the dimension of $\mathscr{H}$ and it tends to infinity as the dimension grows. Halmos believed that M should be a universal constant, independent of T and the dimension of $\mathscr{H}$. In fact, his conjecture is even stronger than believing M to be a universal constant. If (10.4.4) holds, then $w(T^n) \leqslant 1$ and thus Lemma 10.3.7 implies that

$$\|T^n\| \leqslant 2w(T^n) \leqslant 2, \qquad n \geqslant 1.$$

Example 10.4.2 shows that the universal constant 2 is optimal. Brown (unpublished) proved the conjecture for $\dim \mathscr{H} = 2$. Then Bernau and Smithies [8] proved the conjecture for $n = 2^k$. This special case was also independently proved by Fumita, Halmos, and Pearcy (unpublished). Using dilation theory, the conjecture was finally proved by Berger. Shortly after, Pearcy [112] gave an elementary proof of the conjecture.

Berger and Stampfli [7] gave a simplified version of Berger's proof. In fact, they obtained a more general mapping theorem for functions in the disk algebra with $f(0) = 0$. If $w(T) \leqslant 1$, then, for all f in the disk algebra with $f(0) = 0$,

$$w(f(T)) \leqslant \|f\|_\infty.$$

After this period, there was a tremendous amount of research on different types of numerical-range mapping theorems. The one covered in this chapter (Theorem 10.8.4) was discovered about 40 years later by Drury in 2008. He introduced the teardrop region and gave a complete mapping theorem for functions in the disk algebra (not necessarily $f(0) = 0$) [36]. At the heart of the teardrop theorem is an operator inequality, which Drury proved by citing a decomposition theorem of Dritschel and Woerdeman, and then performing some rather complicated calculations. The approach here is adopted from [88] where the authors circumvented these difficulties, and thus simplified Drury's argument, by exploiting finite Blaschke products and a refinement of the inequality in Lemma 10.3.7.

The Wiener Algebra Versus the Disk Algebra Again

The construction in the proof of Theorem 10.5.1 is essentially due to Paul du Bois-Reymond [37]. His goal was to construct a function whose Fourier series diverges at a point of continuity. Working with power series on the disk (rather than a Fourier series) needs some special care. For example, we constructed an $f \in \mathscr{A}(\mathbb{D})$ whose Taylor polynomials do not converge uniformly on $\mathbb{D}^-$. In fact, there is a construction of Sierpinski [127] of a holomorphic function f on $\mathbb{D}$ whose Taylor polynomials converge pointwise on $\mathbb{D}^-$ and yet $f \notin \mathscr{A}(\mathbb{D})$.

10.11 Exercises

10.1 Show that $(\alpha_1 p_1 + \alpha_2 p_2)(T) = \alpha_1 p_1(T) + \alpha_2 p_2(T)$ for all trigonometric polynomials p_1, p_2 and $\alpha_1, \alpha_2 \in \mathbb{C}$.

10.2 Prove that $p(T^*T)T = Tp(TT^*)$ for any analytic polynomial p.

10.3 This exercise outlines another proof of Lemma 10.1.10. Suppose that $T \in \mathscr{L}(\mathscr{H})$ is a contraction and that p is a trigonometric polynomial.

(a) Explain why we may assume that $\|T\| < 1$.
(b) For $\zeta \in \mathbb{T}$, let $S(\zeta) = (I - \overline{\zeta}T)^{-1} + (I - \zeta T^*)^{-1} - I$. Prove that

$$(I - \zeta T^*)S(\zeta)(I - \overline{\zeta}T) = I - T^*T. \tag{10.11.4}$$

(c) Prove that

$$S(\zeta) = (I - \zeta T^*)^{-1}(I - T^*T)(I - \overline{\zeta}T)^{-1}$$

for $\zeta \in \mathbb{T}$. Conclude that $S(\zeta) \geqslant 0$ for all $\zeta \in \mathbb{T}$.
(d) Prove that

$$p(T) = \int_0^{2\pi} p(e^{i\theta})S(e^{i\theta})\frac{d\theta}{2\pi}$$

and conclude that $p(T) \geqslant 0$.

10.4 Show that the operator S defined by (10.1.8) is an isometric dilation of the contraction T.

10.5 Show that the operator U defined by (10.1.9) is a unitary dilation of the isometry S defined by (10.1.8).

10.6 Suppose that $T \in \mathscr{L}(\mathscr{H})$ is a contraction and that p is a trigonometric polynomial with $\|p\|_\infty = 1$.

(a) Show that $a = \operatorname{Re} p$ and $b = \operatorname{Im} p$ are trigonometric polynomials.
(b) Use Lemma 10.1.10 to prove that $-I \leqslant a(T) \leqslant I$ and $-I \leqslant b(T) \leqslant I$.
(c) Conclude that $\|p(T)\| \leqslant 2\|p\|_\infty$.
(d) Apply the preceding result to the n-fold tensor product of T with itself and conclude that $\|p(T)\|^n \leqslant 2\|p\|_\infty^n$.
(e) Deduce von Neumann's inequality (Theorem 10.1.11) from the preceding.

10.7 Prove that the Wiener algebra $\mathscr{W}(\mathbb{D})$ is an algebra. More specifically, show that it is closed under multiplication.

10.8 Provide the details of the proof of Proposition 10.3.2.

10.9 Prove that if $T \in M_n$ is normal, then $W(T)$ is the convex hull of the eigenvalues of T.

10.10 Show that the closure of any convex set is convex.

10.11 Modify the proof of Lemma 10.3.7 to show that $w(T) = \|T\|$ whenever $T \in \mathscr{L}(\mathscr{H})$ is self-adjoint.
Hint: Choose $\gamma \in T$ such that $\gamma\langle Tx, y\rangle = |\langle Tx, y\rangle|$.

10.12 Use 2×2 matrices to show that both of the inequalities in Lemma 10.3.7 can be attained.

10.13 Show that w is not submultiplicative. That is, we do not have the inequality

$$w(TS) \leqslant w(T)w(S), \qquad S, T \in \mathscr{L}(\mathscr{H}).$$

Hint: (Pearcy [112]) Let $\mathscr{H} = \mathbb{C}^4$ and

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Consider $S = N$ and $T = N^2$. One needs to show that

$$w(N^2) = w(N^3) = \tfrac{1}{2} \quad \text{and} \quad w(N) = \tfrac{3}{4}.$$

The last identity might be difficult. However, one can more easily show that $w(N) < 1$, which is enough for this application.

10.14 Show that $\operatorname{td}(\alpha)$ is the intersection of the two families of half planes (10.8.2) and (10.8.3).

10.15 Show that

$$\left|\sum_{k=p}^{q} \frac{e^{ik\theta}}{k}\right| \leqslant \frac{1}{p|\sin(\theta/2)|}.$$

Hint: Use the Abel summation method and the fact that

$$\left|\sum_{k=p}^{q} e^{ik\theta}\right| \leqslant \frac{1}{|\sin(\theta/2)|}.$$

Chapter 11
Real Complex Functions

11.1 Real Rational Functions

In this chapter we connect finite Blaschke products to the class of rational functions f such that

$$f(\zeta) \in \widehat{\mathbb{R}}, \quad \zeta \in \mathbb{T},$$

where we recall that $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ is the extended real line, regarded as a subset of the Riemann sphere $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. We denote the set of such "real rational functions" by $\mathfrak{R}$. Our main focus here is the class

$$\mathfrak{R}^+ := \{f \in \mathfrak{R} : f \text{ is analytic on } \mathbb{D}\}.$$

Example 11.1.1 The function

$$f(z) = i\,\frac{1+z}{1-z}$$

belongs to $\mathfrak{R}^+$ since it is rational, analytic on $\mathbb{D}$, and

$$\begin{aligned} f(e^{i\theta}) &= i\,\frac{1+e^{i\theta}}{1-e^{i\theta}} \\ &= i\,\frac{e^{i\theta/2}(e^{-i\theta/2}+e^{i\theta/2})}{e^{i\theta/2}(e^{-i\theta/2}-e^{i\theta/2})} \\ &= i\,\frac{2\cos(\theta/2)}{-2i\sin(\theta/2)} \\ &= -\cot(\theta/2), \end{aligned}$$

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_11

which belongs to $\widehat{\mathbb{R}}$ for all $\theta \in [0, 2\pi]$. Because f is a Möbius transformation, it maps extended circles in $\widehat{\mathbb{C}}$ to extended circles in $\widehat{\mathbb{C}}$. The values

$$f(-1) = 0, \qquad f(0) = i, \qquad f(1) = \infty, \qquad f(i) = -1, \quad \text{and} \quad f(-i) = 1,$$

show that f maps $\mathbb{D}$ onto the upper-half plane $\mathbb{C}_+$.

Example 11.1.2 Consider the two Blaschke products

$$B_1(z) = \frac{z + \frac{1}{2}}{1 + \frac{1}{2}z} \quad \text{and} \quad B_2(z) = \frac{z - \frac{1}{2}}{1 - \frac{1}{2}z}.$$

A short computation confirms that

$$f(z) = i\,\frac{B_1(z) + B_2(z)}{B_1(z) - B_2(z)} = \frac{3}{2}\frac{iz}{1 - z^2}, \tag{11.1.3}$$

which is rational and analytic on $\mathbb{D}$. Furthermore,

$$\begin{aligned} f(e^{i\theta}) &= \tfrac{3}{2}i\,\frac{e^{i\theta}}{1 - e^{2i\theta}} \\ &= \tfrac{3}{2}i\,\frac{1}{e^{-i\theta} - e^{i\theta}} \\ &= \tfrac{3}{2}i\,\frac{1}{-2i\sin\theta} \\ &= -\tfrac{3}{4}\csc\theta, \end{aligned} \tag{11.1.4}$$

which belongs to $\widehat{\mathbb{R}}$ for all $\theta \in [0, 2\pi]$. Thus, $f \in \mathfrak{R}^+$. In light of (11.1.4), we conclude that

$$f(\mathbb{T}\backslash\{1\}) = (-\infty, -\tfrac{3}{4}] \cup [\tfrac{3}{4}, +\infty)$$

and hence f maps $\mathbb{D}$ onto the complement of the rays $(-\infty, -\frac{3}{4}]$ and $[\frac{3}{4}, +\infty)$. This is illustrated in Fig. 11.1.

Example 11.1.5 Let

$$f(z) = -4\frac{z}{(1 - z)^2}$$

and observe that f is rational and analytic on $\mathbb{D}$. Furthermore,

$$f(e^{i\theta}) = -4\frac{e^{i\theta}}{(1 - e^{i\theta})^2}$$

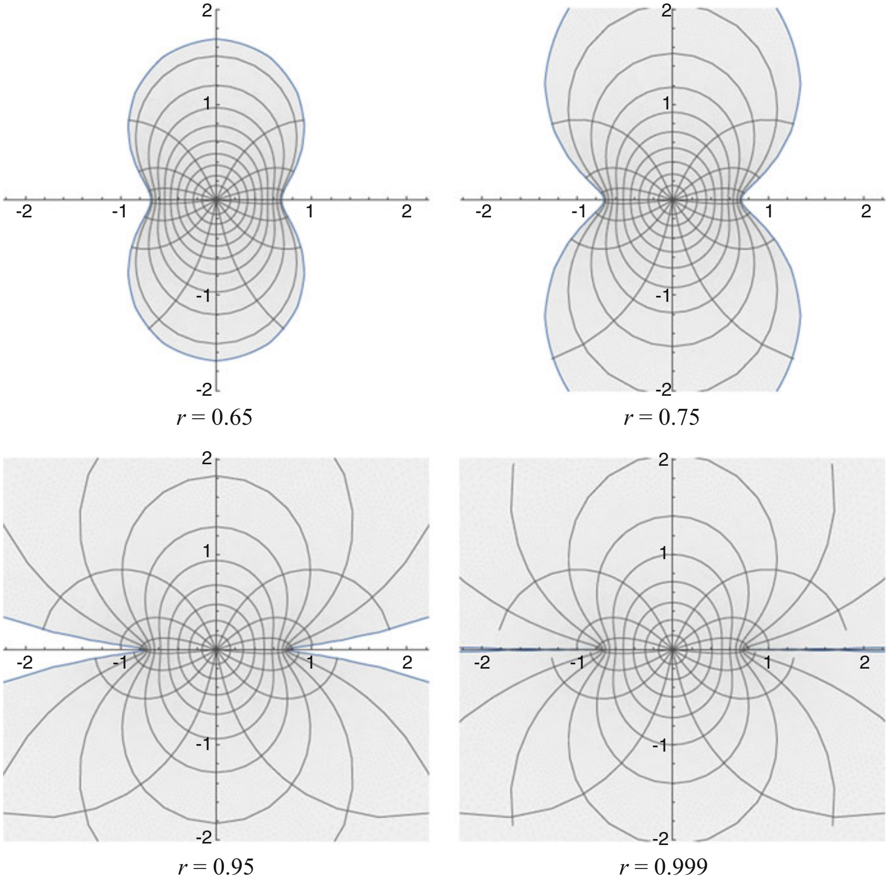


Fig. 11.1 Images of the disk $|z| \leqslant r$ under the function $f(z) = \frac{3}{2}\frac{iz}{1-z^2}$ from (11.1.3) for four values of $r \in (0, 1)$

$$= -4\frac{e^{i\theta}}{(e^{i\theta/2}(e^{-i\theta/2} - e^{i\theta/2}))^2}$$

$$= -4\frac{e^{i\theta}}{e^{i\theta}(-2i\sin(\theta/2))^2}$$

$$= \csc^2(\theta/2), \tag{11.1.6}$$

which belongs to $\widehat{\mathbb{R}}$ for all $\theta \in [0, 2\pi]$. Thus, $f \in \mathfrak{R}^+$. From (11.1.6) we see that

$$f(\mathbb{T}\backslash\{1\}) = [1, \infty)$$

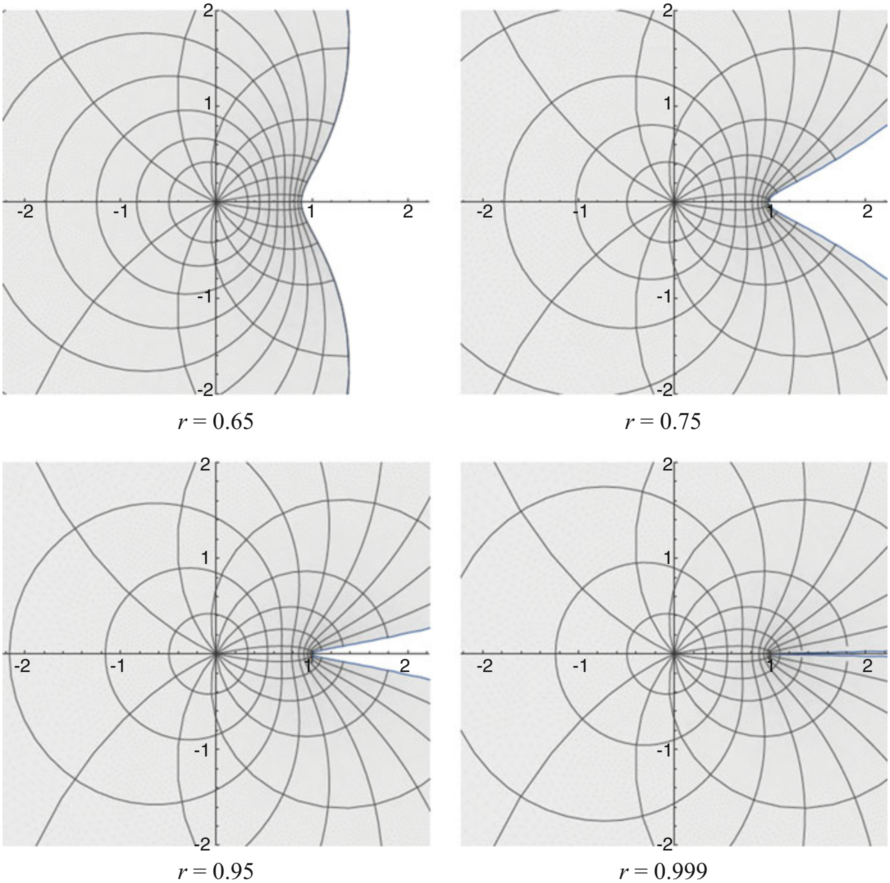


Fig. 11.2 Images of the disk $|z| \leqslant r$ under the function $f(z) = -4z/(1-z)^2$ from Example 11.1.5 for several values of $r \in (0, 1)$

and hence f maps $\mathbb{D}$ onto $\mathbb{C}\backslash[1, \infty)$. This is illustrated in Fig. 11.2. Observe how f is related to the *Koebe function*

$$k(z) = \frac{z}{(1-z)^2},$$

which plays an important role in the study of univalent functions [39, 70, 89].

The alert reader might have noticed that all of the examples of real-rational functions presented above had at least one pole on $\mathbb{T}$. This is no accident.

Proposition 11.1.7 *Suppose f is a nonconstant function in $\mathfrak{R}^+$. Then f has at least one pole on $\mathbb{T}$.*

Proof Suppose toward a contradiction that $f \in \mathfrak{R}^+$ is nonconstant and has no pole on $\mathbb{T}$. Since f is rational and analytic on $\mathbb{D}$, f must actually be analytic in some neighborhood of $\mathbb{D}^-$. The open mapping theorem ensures that $f(\mathbb{D})$ is open. Since f is continuous on $\mathbb{D}^-$, it follows that $f(\mathbb{T})$ is a curve in $\mathbb{C}$ that contains the boundary of $f(\mathbb{D})$. However, f is real valued on $\mathbb{T}$ and hence $f(\mathbb{T})$ is a compact, connected subset of $\mathbb{R}$; that is, $f(\mathbb{T})$ is a closed interval. This forces the open set $f(\mathbb{D})$ to be unbounded, which is a contradiction. □

11.2 Helson's Characterization

For two finite Blaschke products B_1 and B_2, we claim that

$$f = i\frac{B_1 + B_2}{B_1 - B_2}$$

belongs to $\mathfrak{R}$. To see this, observe that

$$f = g\left(\frac{B_1}{B_2}\right),$$

in which

$$g(z) = i\frac{1+z}{1-z} \tag{11.2.1}$$

is the function from Example 11.1.1, which maps $\mathbb{T}$ onto $\widehat{\mathbb{R}}$. Since B_1/B_2 is unimodular on $\mathbb{T}$, this says that $f \in \mathfrak{R}$. A theorem of Helson, presented below, asserts that every function in $\mathfrak{R}$ takes this form.

To state Helson's theorem precisely, we require a definition. Two finite Blaschke products B_1 and B_2 are *relatively prime* if they share no common zeros. Equivalently, there is no nonconstant Blaschke product B for which $B_1 = BC_1$ and $B_2 = BC_2$ for finite Blaschke products C_1, C_2.

Theorem 11.2.2 (Helson [76]) *Suppose f is a rational function.*

(a) If $f \in \mathfrak{R}$, then there are two relatively prime finite Blaschke products B_1 and B_2 such that

$$f = i\frac{B_1 + B_2}{B_1 - B_2}. \tag{11.2.3}$$

(b) If $f \in \mathfrak{R}^+$, then there are two relatively prime finite Blaschke products B_1 and B_2 such that $B_1 - B_2$ has no zeros on $\mathbb{D}$ and (11.2.3) holds.

(c) *If $f \in \mathfrak{R}^+$ has no zeros in $\mathbb{D}$, then there are two relatively prime finite Blaschke products B_1 and B_2 such that $B_1^2 - B_2^2$ has no zeros on $\mathbb{D}$ and (11.2.3) holds.*

Proof

(a) Suppose $f \in \mathfrak{R}$. Let g denote the function (11.2.1) and observe that

$$g^{-1}(z) = \frac{z-i}{z+i}$$

maps $\widehat{\mathbb{R}}$ onto $\mathbb{T}$. It follows that $g^{-1} \circ f$ is a rational function with unimodular boundary values. If f has a pole on $\mathbb{T}$, then $(f-i)/(f+i)$ has the value 1 at this pole. Thus, the preceding quotient is unimodular on $\mathbb{T}$ and meromorphic on $\mathbb{D}$ with a continuous extension to $\mathbb{D}^-$. Corollary 3.5.4 implies that

$$g^{-1} \circ f = \frac{B_1}{B_2} \tag{11.2.4}$$

for two finite Blaschke products B_1 and B_2. By factoring out any common Blaschke factors, we can also assume that B_1 and B_2 are relatively prime. Consequently, we obtain $f = g(B_1/B_2)$, which proves (11.2.3).

(b) Since $\mathfrak{R}^+ \subseteq \mathfrak{R}$, the preceding tells us that f enjoys a representation of the form (11.2.3) with relatively prime finite Blaschke products B_1 and B_2. Suppose toward a contradiction that $B_1 - B_2$, the denominator of f, has a zero $w \in \mathbb{D}$. Since f is analytic on $\mathbb{D}$, the numerator $B_1 + B_2$ must vanish at w as well. Then

$$B_1 + B_2 = b_w G \qquad \text{and} \qquad B_1 - B_2 = b_w H,$$

in which $b_w = (z-w)(1-\overline{w}z)^{-1}$ and G, H are rational analytic functions on $\mathbb{D}$. Solving this system for B_1 and B_2 reveals that

$$B_1 = \tfrac{1}{2} b_w (G+H) \qquad \text{and} \qquad B_2 = \tfrac{1}{2} b_w (G-H).$$

Thus, B_1 and B_2 have a common zero at w. This is a contradiction to the fact that B_1 and B_2 were chosen to be relatively prime.

(c) Proceeding as in (b) we see that $B_1 + B_2$ also has no zeros on $\mathbb{D}$. Thus,

$$B_1^2 - B_2^2 = (B_1 + B_2)(B_1 - B_2)$$

has no zeros on $\mathbb{D}$ as well. □

To obtain the Helson decomposition for a general $f \in \mathfrak{R}$, use the identity

$$\frac{f(z)-i}{f(z)+i} = \frac{B_1(z)}{B_2(z)}$$

and compute

$$\{w_1, w_2, \ldots, w_n\} = f^{-1}(\{i\}) \cap \mathbb{D},$$

and

$$\{\lambda_1, \lambda_2 \ldots, \lambda_m\} = f^{-1}(\{-i\}) \cap \mathbb{D}.$$

Then

$$B_1(z) = \xi \prod_{j=1}^{n} \frac{z - w_j}{1 - \overline{w_j} z} \qquad \text{and} \qquad B_2(z) = \prod_{j=1}^{m} \frac{z - \lambda_j}{1 - \overline{\lambda_j} z},$$

for some unimodular constant ξ.

Example 11.2.5 Let us compute the Helson decomposition of the function

$$f(z) = -4\frac{z}{(1-z)^2}$$

from Example 11.1.5. From (11.2.4),

$$\frac{f(z) - i}{f(z) + i} = -\frac{z^2 - 2(1+2i)z + i}{z^2 - 2(1-2i)z + i} = \frac{B_1}{B_2}.$$

The single zero of B_1 is the solution to $f(z) = i$, or

$$z^2 - 2(1+2i)z + i = 0,$$

that belongs to $\mathbb{D}$. A computation shows that

$$w \approx 0.0898203 - 0.197368i.$$

Similarly, the single zero of B_2 is the solution to $f(z) = -i$ that lies in $\mathbb{D}$, which turns out to be $\overline{w}$. The unimodular constant factors in B_1 and B_2 should be chosen so that

$$-1 = \frac{f(0) - i}{f(0) + i} = \frac{B_1(0)}{B_2(0)}.$$

Thus,

$$B_1(z) = \xi \frac{z - w}{1 - \overline{w}z} \qquad \text{and} \qquad B_2(z) = \frac{z - \overline{w}}{1 - wz},$$

in which $\xi = -w/\overline{w}$. This yields the Helson representation

$$f(z) = i\,\frac{B_1(z) + B_2(z)}{B_1(z) - B_2(z)}.$$

11.3 Real Rational Functions Without Zeros

Theorem 11.2.2 says that $f \in \mathfrak{R}^+$ has no zeros on $\mathbb{D}$ if and only if

$$f = i\,\frac{B_1 + B_2}{B_1 - B_2}$$

for two relatively prime finite Blaschke products B_1 and B_2 so that $B_1^2 - B_2^2$ has no zeros on $\mathbb{D}$. Our aim in this section is to obtain a more precise description of these nonvanishing $\mathfrak{R}^+$ functions.

Example 11.3.1 For $e^{i\alpha}, e^{i\beta} \in \mathbb{T}$, consider the function

$$f_{\alpha,\beta}(z) := e^{-i(\frac{\alpha-\beta}{2})}\left(\frac{e^{i\alpha} - z}{e^{i\beta} - z}\right). \tag{11.3.2}$$

Observe that $f_{\alpha,\beta}$ is rational, analytic on $\mathbb{D}$, and satisfies

$$\begin{aligned} f_{\alpha,\beta}(e^{i\theta}) &= \frac{e^{i\alpha/2} - e^{-i\alpha/2}e^{i\theta}}{e^{i\beta/2} - e^{-i\beta/2}e^{i\theta}} \\ &= \frac{e^{i(\alpha-\theta)/2} - e^{-i(\alpha-\theta)/2}}{e^{i(\beta-\theta)/2} - e^{-i(\beta-\theta)/2}} \\ &= \left[\sin\left(\frac{\theta-\alpha}{2}\right)/\sin\left(\frac{\theta-\beta}{2}\right)\right]. \end{aligned}$$

This means that $f_{\alpha,\beta}$ belongs to $\mathfrak{R}^+$ and has no zeros on $\mathbb{D}$. For example, letting $\alpha = \beta + \pi$ yields the function

$$i\,\frac{\xi + z}{\xi - z}, \tag{11.3.3}$$

which belongs to $\mathfrak{R}^+$.

We now prove that these functions $f_{\alpha,\beta}$ are the building blocks for $\mathfrak{R}^+$ functions without zeros in $\mathbb{D}$.

Theorem 11.3.4 *Suppose $f \in \mathfrak{R}^+$ and has no zeros on $\mathbb{D}$. Then either f is a nonzero real constant function or*

$$f(z) = c\xi \prod_{j=1}^{n} \frac{e^{i\alpha_j} - z}{e^{i\beta_j} - z},$$

for some $c \in \mathbb{R}$, $\xi \in \mathbb{T}$, *and* $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n \in [0, 2\pi]$.

Proof Theorem 11.2.2 says that

$$f = i\frac{B_1 + B_2}{B_1 - B_2},$$

where B_1 and B_2 are relatively prime finite Blaschke products such that $B_1 - B_2$ and $B_1 + B_2$ have no zeros on $\mathbb{D}$. Consequently, any zeros or poles of f must lie in $|z| \geqslant 1$. However, the reflection identity (3.1.6) implies that the zeros or poles of f cannot be in $|z| > 1$. Thus, any zeros or poles of f belong to $\mathbb{T}$.

The zeros of f occur at those z for which

$$\frac{B_1(z)}{B_2(z)} = -1.$$

Clear the denominators in the preceding and rewrite it as a polynomial equation in z. Suppose that the resulting equation has $n \geqslant 0$ solutions

$$e^{i\alpha_1}, e^{i\alpha_2}, \ldots, e^{i\alpha_n}.$$

Similarly, the poles occur at those z for which

$$\frac{B_1(z)}{B_2(z)} = 1.$$

Suppose that this equation has $m \geqslant 0$ solutions

$$e^{i\beta_1}, e^{i\beta_2}, \ldots, e^{i\beta_m}.$$

Since f is a rational function with no other zeros or poles, we must have

$$f(z) = c\xi \left(\prod_{j=1}^{n} (e^{i\alpha_j} - z) \right) \prod_{j=1}^{m} \frac{1}{e^{i\beta_j} - z},$$

for some real constant c and some unimodular constant ξ. There are several possibilities.

(a) If $n = m = 0$, then f is a real constant function.
(b) If $m = 0$ and $n > 0$, then

$$f(z) = c\xi \prod_{j=1}^{n} (e^{i\alpha_j} - z),$$

which says that $f \in \mathfrak{R}^+$ is nonconstant and bounded. However, this contradicts Proposition 11.1.7.

(c) If $n = 0$ and $m > 0$, then

$$\frac{1}{f} = c^{-1}\overline{\xi} \prod_{j=1}^{n} (e^{i\beta_j} - z),$$

which implies that $1/f \in \mathfrak{R}^+$ is nonconstant and bounded. This also contradicts Proposition 11.1.7.

(d) If $m > 0$, $n > 0$, but $m < n$, then for some appropriately chosen unimodular constant ζ, the function

$$g = \zeta \prod_{j=1}^{m} \frac{e^{i\alpha_j} - z}{e^{i\beta_j} - z},$$

along with f/g, belongs to $\mathfrak{R}^+$. Then

$$\frac{f}{g} = \overline{\zeta} c\xi \prod_{j=m+1}^{n} (e^{i\alpha_j} - z)$$

belongs to $\mathfrak{R}^+$ and is bounded (and nonconstant). However, this contradicts Proposition 11.1.7.

(e) A similar contradiction arises when $m > 0$, $n > 0$, but $n < m$.

Thus, $m = n$ and we have the desired factorization

$$f(z) = c\xi \prod_{j=1}^{n} \frac{e^{i\alpha_j} - z}{e^{i\beta_j} - z}$$

for some real constant c and some unimodular constant ξ. □

11.4 Factorization

One can factor any analytic function f on $\mathbb{D}$ as $f = FG$, where F is analytic on $\mathbb{D}$ and whose zeros are precisely those of f (including multiplicity) and G is analytic on $\mathbb{D}$ with no zeros on $\mathbb{D}$ [26, Thm. 5.25]. When f belongs to a certain class of functions, one often wants a factorization $f = FG$ in which F and G not only satisfy the properties above but also belong to the same class of functions as f. For example, if f is a bounded analytic function on $\mathbb{D}$, then $f = BG$, where B is a bounded analytic function on $\mathbb{D}$ with the same zeros of f (including multiplicity) and G is a bounded analytic function on $\mathbb{D}$ with no zeros [38].

Suppose that $f \in \mathfrak{R}^+$. Then one can write $f = BG$, where B is a finite Blaschke product whose zeros are precisely those of f, repeated according to multiplicity, and $G = f/B$. However, it will not always be the case that the two factors belong to $\mathfrak{R}^+$. The following theorem remedies this situation.

Theorem 11.4.1 *If $f \in \mathfrak{R}^+$, then $f = FG$, in which*

(a) $F, G \in \mathfrak{R}^+$,
(b) F has precisely the same zeros of f (with the same multiplicities),
(c) G has no zeros in $\mathbb{D}$,
(d) $|f| \leqslant |G|$ on $\mathbb{T}$,
(e) f and G have the same sign on $\mathbb{T}$.

Proof Let $f \in \mathfrak{R}^+$. Since f is a rational function, it has a finite number of zeros in $\mathbb{D}$. If f has no zeros in $\mathbb{D}$, then let $B \equiv i$. Otherwise, let B be a finite Blaschke product whose zeros are those of f, with the same multiplicities. Set $H = f/B$ and observe that H has no zeros on $\mathbb{D}$. Furthermore,

$$f = \frac{-4B}{(1-B)^2} \cdot \frac{(1-B)^2}{-4} H = FG,$$

where

$$F = \frac{-4B}{(1-B)^2} \quad \text{and} \quad G = \frac{(1-B)^2}{-4} H.$$

If f has no zeros in $\mathbb{D}$, then F is a real constant function and G is a rational function. From Example 11.1.5, observe that $F \in \mathfrak{R}^+$ and $F \geqslant 1$ on $\mathbb{T}$. This says that $G \in \mathfrak{R}^+$ (and has no zeros on $\mathbb{D}$ by construction) and $|G| = |f/F| \leqslant |f|$ on $\mathbb{T}$. Finally, since $f/G = F \geqslant 1$ on $\mathbb{T}$, we see that f and G have the same sign on $\mathbb{T}$. □

11.5 Valence

Recall that $|E|$ denotes the cardinality of a set E. For $f \in \mathfrak{R}^+$ and $w \in \mathbb{C}\backslash\mathbb{R}$, let

$$v(f, w) = |\{z \in \mathbb{D} : f(z) = w\}|$$

denote the *valence* of f at w. This is the number of times, counting multiplicity, that f assumes the value w in $\mathbb{D}$. By Theorem 11.2.2, each $f \in \mathfrak{R}^+$ takes the form

$$f = i\frac{B_1 + B_2}{B_1 - B_2},$$

in which B_1, B_2 are finite Blaschke products with no common zeros and such that $B_1 - B_2$ has no zeros on $\mathbb{D}$. Since finite Blaschke products have constant valence on $\mathbb{D}$ and $\mathbb{C}\backslash\mathbb{D}^-$ (Theorem 3.4.10), this should translate into information about the valence of f on $\mathbb{C}_+$ and $\mathbb{C}_-$.

Theorem 11.5.1 *Let $f \in \mathfrak{R}^+$ be of the form*

$$f = i\frac{B_1 + B_2}{B_1 - B_2},$$

where B_1, B_2 are finite Blaschke products with no common zeros and such that $B_1 - B_2$ has no zeros on $\mathbb{D}$. Then

$$v(f, w) = \begin{cases} \deg B_2 & \text{if } w \in \mathbb{C}_+, \\ \deg B_1 & \text{if } w \in \mathbb{C}_-. \end{cases}$$

In other words, f has constant valence on each of $\mathbb{C}_+$ and $\mathbb{C}_-$.

Proof First observe that the Möbius transformation

$$\psi(z) = \frac{z - i}{z + i}$$

is injective; it maps $\mathbb{C}_+$ onto $\mathbb{D}$ and it maps $\mathbb{C}_-$ onto $\widehat{\mathbb{C}}\backslash\mathbb{D}^-$. There are several cases to consider.

(a) If $w \in \mathbb{C}_+$, then the number of solutions to $f(z) = w$ is the same as the number of solutions to

$$\psi \circ f(z) = \psi(w) = \eta \in \mathbb{D}.$$

This, in turn, equals the number of solutions in $\mathbb{D}$ to

$$\eta = \frac{f(z) - i}{f(z) + i} = \frac{B_2(z)}{B_1(z)}.$$

We need to examine the number of zeros of $B_2 - \eta B_1$ in $\mathbb{D}$. Since

$$|\eta B_1| = |\eta| < 1 = |B_2|$$

on $\mathbb{T}$, Rouché's theorem (see Sect. 4.5) implies that the number of zeros of B_2 and $B_2 - \eta B_1$ in $\mathbb{D}$ are the same. Thus, $v(f, w) = \deg B_2$ whenever $w \in \mathbb{C}_+$.

(b) If $w \in \mathbb{C}_-$ and $w \neq -i$, then $\eta = \psi(w) \in \mathbb{C}\backslash\mathbb{D}^-$. In particular, $\eta = \psi(w) \neq \infty$. We want to count the number of zeros of $B_2 - \eta B_1$ in $\mathbb{D}$. Since

$$|B_2| = 1 < |\eta| = |\eta B_1|$$

on $\mathbb{T}$, Rouché's theorem implies that ηB_1 and $B_2 - \eta B_1$ have the same number of zeros in $\mathbb{D}$. Thus, $v(f, w) = \deg B_1$ for any $w \in \mathbb{C}_-\backslash\{-i\}$.

(c) If $w = -i$, then we need to find the number of solutions in $\mathbb{D}$ to

$$\frac{B_2}{B_1} = \infty.$$

Since B_1 and B_2 are relatively prime, this is the same as the number of zeros of B_1, which is $\deg B_1$.

This completes the proof. □

Theorem 11.5.1 shows that any $f \in \mathfrak{R}^+$ has constant valence on each of $\mathbb{C}_+$ and $\mathbb{C}_-$. It turns out that we can make these two (constant) valences anything we want.

Theorem 11.5.2 (Helson [75]) *For a given pair of $m, n \in \mathbb{N} \cup \{0\}$, there is an $f \in \mathfrak{R}^+$ with valence m on $\mathbb{C}_+$ and valence n on $\mathbb{C}_-$.*

Proof Fix $m, n \in \mathbb{N} \cup \{0\}$ and let $P = \{\zeta_1, \zeta_2, \ldots, \zeta_m\}$ and $N = \{\xi_1, \xi_2, \ldots, \xi_n\}$ be sets of distinct points on $\mathbb{T}$ with $P \cap N = \varnothing$. If either m or n equals zero, take the corresponding set P or N to be the empty set. If $m = n = 0$ set $f \equiv 1$ and note that $f \in \mathfrak{R}^+$ with $v_f \equiv 0$ of both $\mathbb{C}_+$ and $\mathbb{C}_-$. If $m, n \in \mathbb{N}$, define

$$f(z) = i\sum_{j=1}^{m} \frac{\zeta_j + z}{\zeta_j - z} - i\sum_{k=1}^{n} \frac{\xi_k + z}{\xi_k - z}$$

and observe that f is a rational function which belongs to $\mathfrak{R}^+$ in light of (11.3.3). Notice that $f(\zeta_j) = \infty$ and that f maps a neighborhood of ζ_j in $\mathbb{D}$ to a neighborhood of ∞ in $\mathbb{C}_+$. Since $\zeta_1, \zeta_2, \ldots, \zeta_m$ are distinct, f maps $\mathbb{D}$ onto a neighborhood of ∞ in $\mathbb{C}_+$ exactly m times. By Theorem 11.5.1, the valence of f equals m on all of $\mathbb{C}_+$. In a similar way, $f(\xi_k) = \infty$ and f maps a neighborhood of ξ_k in $\mathbb{D}$ to a neighborhood of ∞ in $\mathbb{C}_-$. Since $\xi_1, \xi_2, \ldots, \xi_n$ are distinct, f maps $\mathbb{D}$ onto a neighborhood of ∞ in $\mathbb{C}_-$ exactly n times. Thus, f has valence n on all of $\mathbb{C}_-$. The arguments above even take care of the case when either m or n are equal to zero. □

The observant reader might want to use Theorem 11.5.1 to produce an $f \in \mathfrak{R}^+$ with prescribed valences. Indeed, in the representation

$$f = i\frac{B_1 + B_2}{B_1 - B_2},$$

one just needs to make the degrees of B_1 and B_2 match the desired valences on $\mathbb{C}_+$ and $\mathbb{C}_-$. However, in order for f to belong to $\mathfrak{R}^+$, we also need $B_1 - B_2$ to have no zeros in $\mathbb{D}$, which is not always easy to do.

11.6 Notes

The study of “real complex" functions goes beyond the rational setting we presented in this chapter [56, 60, 75, 76]. In this more general setting, finite Blaschke products are replaced by inner functions.

Connection with Model Spaces

There is a connection between real rational functions and model spaces with kernels of Toeplitz operators [51, 53].

11.7 Exercises

11.1 Explore the Helson decomposition of the Koebe function

$$k(z) = \frac{z}{(1-z)^2}$$

and for $k(z)^2$.

11.2 Compute the Helson decomposition of the function $f_{\alpha,\beta}(z)$ from (11.3.2).

11.3 For $f \in \mathfrak{R}^+$, what types of domains can be the image $f(\mathbb{D})$ of f?

11.4 Consider the Möbius transformation

$$T(z) = i\,\frac{1-iz}{1+iz}$$

and observe that T maps $\mathbb{D}$ onto $\mathbb{C}_+$ and maps $\mathbb{T}$ onto $\widehat{\mathbb{R}}$. Prove the following identities.

(a) $T^{-1}(z) = T(1/z) = \dfrac{1}{T(z)} = -T(-z) = \overline{T(\overline{z})}$.

(b) $(T \circ T)(z) = \dfrac{1}{z}$ and $(T \circ T \circ T \circ T)(z) = z$.

(c) $T(z_1 z_2) = \dfrac{T(z_1)T(z_2) + T(z_1) + T(z_2) - 1}{1 + T(z_1) + T(z_2) - T(z_1)T(z_2)}$.

(d) $T(z_2/z_1) = \dfrac{T(z_1)T(z_2) - T(z_1) + T(z_2) + 1}{T(z_1)T(z_2) + T(z_1) - T(z_2) + 1}$.

(e) $T(z_1 + z_2) = \dfrac{3T(z_1)T(z_2) + iT(z_1) + iT(z_2) + 1}{3i + T(z_1) + T(z_2) + iT(z_1)T(z_2)}$.

11.5 Suppose that $f_1 = T(B_1)$ and $f_2 = T(B_2)$ for some finite Blaschke products B_1 and B_2. Then $f_1 + f_2$ is rational, real valued on $\mathbb{R}$, and maps $\mathbb{D}$ into $\mathbb{C}$. Consequently, Exercise 11.4 and a rephrasing of Helson's theorem provide a finite Blaschke product B so that $f_1 + f_2 = T(B)$. Use the identities in Exercise 11.4 to show that

$$B = \frac{3iB_1B_2 + B_1 + B_2 + i}{3 + iB_1 + iB_2 + B_1B_2}.$$

11.6 Helson's theorem tells us that each $f \in \mathfrak{R}$ can be written as $f = T(B_2/B_1)$, where B_1 and B_2 are relatively prime finite Blaschke products. Using the identities in Exercise 11.4, show that if $f_1 = T(B_1)$ and $f_2 = T(B_2)$, then

$$f = \frac{f_1f_2 - f_1 + f_2 + 1}{f_1f_2 + f_1 - f_2 + 1}.$$

Chapter 12
Finite-Dimensional Model Spaces

Model spaces are of great importance to operator theory since the corresponding compressions of the shift operator (see (12.6.4) below) can be used to represent certain contractions [59, 106, 131]. In this chapter, we develop finite-dimensional model spaces using elementary techniques to give the reader a taste of a much larger picture.

The reader familiar with model spaces in their full glory may complain that we are not as comprehensive as we should be. For the sake of the novice, however, our presentation involves little more than the Cauchy integral formula (in the guise of Lemma A.2.2 below) and linear algebra. The interested reader can further explore model spaces and their applications in the recent text [59].

12.1 Model Spaces

Definition 12.1.1 For a finite Blaschke product B with zeros $\lambda_1, \lambda_2, \ldots, \lambda_n$, repeated according to multiplicity, the *model space* $\mathscr{K}_B$ is the set of rational functions of the form

$$\mathscr{K}_B := \left\{ \frac{P(z)}{R(z)} : P \in \mathscr{P}_{n-1} \right\}, \tag{12.1.2}$$

in which

$$R(z) = (1 - \overline{\lambda_1} z)(1 - \overline{\lambda_2} z) \cdots (1 - \overline{\lambda_n} z). \tag{12.1.3}$$

In what follows, R denotes the denominator (12.1.3) that appears in the definition of $\mathscr{K}_B$. Its degree and the location of its zeros can be inferred from context. Since

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_12

$\mathscr{P}_{n-1}$ is the set of (analytic) polynomials of degree at most $n-1$, it follows that $\mathscr{K}_B$ is a complex vector space and

$$\dim \mathscr{K}_B = n. \tag{12.1.4}$$

Example 12.1.5 If $B(z) = z^n$, then $R(z) \equiv 1$ and hence $\mathscr{K}_{z^n} = \mathscr{P}_{n-1}$.

Example 12.1.6 If $\lambda \in \mathbb{D}$ and

$$B(z) = \frac{z-\lambda}{1-\overline{\lambda}z},$$

then $R(z) = 1 - \overline{\lambda}z$ and $n = 1$. Thus,

$$\mathscr{K}_B = \operatorname{span}\left\{\frac{1}{1-\overline{\lambda}z}\right\}.$$

Example 12.1.7 If $\lambda \in \mathbb{D}$ and

$$B(z) = \left(\frac{z-\lambda}{1-\overline{\lambda}z}\right)^2,$$

then $R(z) = (1-\overline{\lambda}z)^2$ and $n = 2$. Partial fractions confirm that

$$\mathscr{K}_B = \operatorname{span}\left\{\frac{1}{1-\overline{\lambda}z}, \frac{1}{(1-\overline{\lambda}z)^2}\right\}.$$

Example 12.1.8 If $\lambda, \eta \in \mathbb{D}$ with $\lambda \neq \eta$ and

$$B(z) = \frac{z-\lambda}{1-\overline{\lambda}z} \cdot \frac{z-\eta}{1-\overline{\eta}z},$$

then $R(z) = (1-\overline{\lambda}z)(1-\overline{\eta}z)$ and $n = 2$. Consequently,

$$\mathscr{K}_B = \operatorname{span}\left\{\frac{1}{1-\overline{\lambda}z}, \frac{1}{1-\overline{\eta}z}\right\}.$$

A useful observation is the following.

Proposition 12.1.9 *$\mathscr{K}_B$ contains the constant functions if and only if $B(0) = 0$.*

Proof Let B denote a finite Blaschke product of degree n with zeros $\lambda_1, \lambda_2, \ldots, \lambda_n$, repeated according to multiplicity. If $B(0) = 0$, then $\deg R \leqslant n-1$. Thus, $P = cR \in \mathscr{P}_{n-1}$ for all $c \in \mathbb{C}$ and hence $c = cR/R \in \mathscr{K}_B$. Conversely, suppose that $c \in \mathscr{K}_B$ for some $c \in \mathbb{C}\backslash\{0\}$. Then $cR = P$ for some $P \in \mathscr{P}_{n-1}$. In light of (12.1.3), this can occur if and only if $\deg R \leqslant n-1$; that is, if at least one of $\lambda_1, \lambda_2, \ldots, \lambda_n$ is zero. Thus, $B(0) = 0$. □

We can make $\mathscr{K}_B$ into a Hilbert space by endowing it with the inner product

$$\langle f, g\rangle = \int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})}\,\frac{d\theta}{2\pi} \tag{12.1.10}$$

and norm

$$\|f\| = \sqrt{\langle f, f\rangle} = \left(\int_0^{2\pi} |f(e^{i\theta})|^2\,\frac{d\theta}{2\pi}\right)^{1/2}$$

it inherits by regarding the elements of $\mathscr{K}_B$ as members of the Lebesgue space $L^2 := L^2(\mathbb{T})$ (see the Appendix). That is, we can identify $f \in \mathscr{K}_B$ with its boundary function $f : \mathbb{T} \to \mathbb{C}$ and consider $\mathscr{K}_B$ as a finite-dimensional subspace of L^2. This permits us to refer to the inner product (12.1.10) between any two rational functions with no poles in $\mathbb{D}^-$. Indeed, the boundary functions of such rational functions are square-integrable on $\mathbb{T}$ and hence (12.1.10) is well defined for such f, g.

Building upon Example 12.1.6, we identify some conspicuous residents of $\mathscr{K}_B$. For $\lambda \in \mathbb{D}$, the corresponding *Cauchy kernel* is

$$c_\lambda(z) = \frac{1}{1-\overline{\lambda}z}. \tag{12.1.11}$$

The Cauchy integral formula (Lemma A.2.2) implies that

$$\langle f, c_\lambda\rangle = f(\lambda), \qquad f \in \mathscr{K}_B. \tag{12.1.12}$$

More generally, this holds for any rational function with no poles in $\mathbb{D}^-$ if we regard the inner product as being performed in L^2.

Example 12.1.13 We can use (12.1.12) to compute $\|c_\lambda\|$:

$$\begin{aligned}\|c_\lambda\|^2 &= \langle c_\lambda, c_\lambda\rangle \\ &= c_\lambda(\lambda) \\ &= \frac{1}{1-|\lambda|^2}.\end{aligned}$$

In what follows, it is sometimes convenient to use the normalized version of c_λ:

$$\widetilde{c}_\lambda = \frac{\sqrt{1-|\lambda|^2}}{1-\overline{\lambda}z}. \tag{12.1.14}$$

By construction, $\widetilde{c}_\lambda$ is a unit vector.

A computation similar to (12.1.12) reveals that

$$\langle f, c_\lambda^{(j)}\rangle = f^{(j)}(\lambda),$$

in which

$$c_\lambda^{(j)} = \frac{j!z^j}{(1-\overline{\lambda}z)^{j+1}} \tag{12.1.15}$$

is the jth derivative, with respect to $\overline{\lambda}$, of the Cauchy kernel $c_\lambda(z)$; see Exercise 12.1. The functions (12.1.11) and (12.1.15) appear frequently in what follows.

Proposition 12.1.16 *If B is a finite Blaschke product with distinct zeros $\lambda_1, \lambda_2, \ldots, \lambda_r$ and corresponding multiplicities $m_1, m_2, \ldots, m_r$, then*

$$\mathscr{K}_B = \operatorname{span}\left\{c_{\lambda_i}^{(j)} : 1 \leqslant i \leqslant r,\ 0 \leqslant j \leqslant m_i - 1\right\}.$$

Proof We prove this under the assumption that each of the zeros of B is simple; that is, $m_1 = m_2 = \cdots = m_r = 1$. See Exercise 12.2 for the general case.

If $P/R \in \mathscr{K}_B$, in which $P \in \mathscr{P}_{n-1}$ and R is given by (12.1.3), then a partial fraction expansion renders P/Q as a linear combination of $c_{\lambda_1}, c_{\lambda_2}, \ldots, c_{\lambda_n}$.

Conversely, any linear combination $\sum_{j=1}^n a_j c_{\lambda_j}$ can be put over the common denominator R by writing

$$\sum_{j=1}^n a_j c_{\lambda_j}(z) = \frac{\sum_{j=1}^n a_j R(z)/(1-\overline{\lambda_j}z)}{R(z)} = \frac{P(z)}{R(z)},$$

in which

$$\deg P \leqslant \deg R - 1 \leqslant n - 1.$$

Thus, any such linear combination belongs to $\mathscr{K}_B$. □

One can see from (12.1.2) that the elements of $\mathscr{K}_B$ are rational functions with no poles in $\mathbb{D}^-$. Using the ideas above, we can give the following characterization of $\mathscr{K}_B$ in terms of inner products.

Proposition 12.1.17 *A rational function f with no poles in $\mathbb{D}^-$ belongs to $\mathscr{K}_B$ if and only if $\langle f, Bz^k\rangle = 0$ for all $k \geqslant 0$.*

Proof We prove this under the assumption that each of the zeros $\lambda_1, \lambda_2, \ldots, \lambda_n$ of B is simple; see Exercise 12.3 for the general case. Under this assumption, Proposition 12.1.16 implies that

$$\mathscr{K}_B = \operatorname{span}\{c_{\lambda_1}, c_{\lambda_2}, \ldots, c_{\lambda_n}\}.$$

($\Rightarrow$) If $f \in \mathscr{K}_B$, then

$$f = \sum_{j=1}^n a_j c_{\lambda_j}$$

for some $a_1, a_2, \ldots, a_n \in \mathbb{C}$. For any $k \geqslant 0$,

$$\begin{aligned}\langle z^k B, f\rangle &= \Big\langle z^k B, \sum_{j=1}^{n} a_j c_{\lambda_j}\Big\rangle \\ &= \sum_{j=1}^{n} \overline{a_j}\langle z^k B, c_{\lambda_j}\rangle \\ &= \sum_{j=1}^{n} \overline{a_j}\lambda_j^k B(\lambda_j) \\ &= 0.\end{aligned}$$

($\Leftarrow$) Suppose that f is a rational function with no poles in $\mathbb{D}^-$ and $\langle z^k B, f\rangle = 0$ for $k \geqslant 0$. We assume that the poles of f are finite and simple; for the general case, see Exercise 12.3. Then we may write

$$f = \sum_{j=1}^{m} a_j c_{w_j}, \tag{12.1.18}$$

in which $w_1, w_2, \ldots, w_m \in \mathbb{D}$ are distinct and $a_1, a_2, \ldots, a_m$ are nonzero. For $i \in \{1, 2, \ldots, m\}$, use the Lagrange interpolation theorem (Theorem 7.1.1) to obtain polynomials p_i so that $p_i(w_j) = \delta_{i,j}$. Then (12.1.12) yields

$$\begin{aligned}0 &= \langle p_i B, f\rangle \\ &= \langle p_i B, \sum_{j=1}^{m} a_j c_{w_j}\rangle \\ &= \sum_{j=1}^{m} \overline{a_j}\langle p_i B, c_{w_j}\rangle \\ &= \sum_{j=1}^{m} a_j p_i(w_j) B(w_j) \\ &= a_i B(w_i).\end{aligned}$$

Thus, $B(w_i) = 0$ and hence w_i is a zero of B; that is, $w_i = \lambda_j$ for some j. Consequently,

$$f \in \operatorname{span}\{c_{\lambda_1}, c_{\lambda_2}, \ldots, c_{\lambda_n}\} = \mathscr{K}_B.$$

This completes the proof. □

If $f \in \mathcal{K}_{\mathcal{B}}$, then we may write

$$f(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \cdots,$$

in which the a_n are the Taylor coefficients at the origin for f. The *backward shift* of f is the function

$$\frac{f(z) - f(0)}{z} = a_1 + a_2 z^2 + a_3 z^2 + a_4 z^3 + \cdots. \tag{12.1.19}$$

The following corollary shows that finite-dimensional model spaces are invariant under the backward shift.

Corollary 12.1.20 *If $f \in \mathcal{K}_B$, then*

$$\frac{f(z) - f(0)}{z} \in \mathcal{K}_B.$$

Proof If $f \in \mathcal{K}_B$ and $k \geqslant 0$, then

$$\begin{aligned}\left\langle \frac{f - f(0)}{z}, z^k B \right\rangle &= \langle f - f(0), z^{k+1} B \rangle \\ &= \langle f, z^{k+1} B \rangle - f(0) \langle 1, z^{k+1} B \rangle.\end{aligned}$$

The first inner product is zero by Proposition 12.1.17. The second inner product is zero by Cauchy's Theorem. Proposition 12.1.17 implies that $(f - f(0))/z \in \mathcal{K}_B$. □

12.2 The Takenaka Basis

The computation

$$\begin{aligned}\langle z^j, z^k \rangle &= \int_0^{2\pi} e^{ij\theta} \overline{e^{ik\theta}} \, \frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} e^{i(j-k)\theta} \, \frac{d\theta}{2\pi} \\ &= \delta_{j,k}\end{aligned}$$

shows that $\{1, z, z^2, \ldots, z^{n-1}\}$ is an orthonormal basis for

$$\mathcal{K}_{z^n} = \mathcal{P}_{n-1} = \operatorname{span}\{1, z, z^2, \ldots, z^{n-1}\}.$$

One construction of an orthonormal basis for a general finite-dimensional model space $\mathscr{K}_B$ is due to Takenaka [132]. To simplify things, we employ the shorthand

$$b_\lambda(z) = \frac{z-\lambda}{1-\overline{\lambda}z} \tag{12.2.1}$$

for $\lambda \in \mathbb{D}$.

Proposition 12.2.2 *Let B be a finite Blaschke product with zeros $\lambda_1, \lambda_2, \ldots, \lambda_n$, repeated according to multiplicity. Let $v_1 = \widetilde{c}_{\lambda_1}$ and*

$$v_\ell = (b_{\lambda_1} \cdots b_{\lambda_{\ell-1}})\widetilde{c}_{\lambda_\ell}, \tag{12.2.3}$$

for $2 \leqslant \ell \leqslant n$. Then $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal basis for $\mathscr{K}_B$.

Proof Since $b_{\lambda_i}\overline{b_{\lambda_i}} = 1$ on $\mathbb{T}$ for $i = 1, 2, \ldots, n$, the definition (12.1.10) of the inner product on $\mathscr{K}_B$ ensures that for $j < k$ we have

$$\begin{aligned}
\langle v_k, v_j\rangle &= \langle b_{\lambda_1}b_{\lambda_2}\cdots b_{\lambda_{k-1}}\widetilde{c}_{\lambda_k},\ b_{\lambda_1}b_{\lambda_2}\cdots b_{\lambda_{j-1}}\widetilde{c}_{\lambda_j}\rangle \\
&= \sqrt{1-|a_j|^2}\langle b_{\lambda_j}b_{\lambda_{j+1}}\cdots b_{\lambda_{k-1}}\widetilde{c}_{\lambda_k},\ c_{\lambda_j}\rangle \\
&= \sqrt{1-|a_j|^2}b_{\lambda_j}(\lambda_j)b_{\lambda_{j+1}}(\lambda_j)\cdots b_{\lambda_{k-1}}(\lambda_j)\widetilde{c}_{\lambda_k}(\lambda_j) \\
&= 0
\end{aligned}$$

since $b_{\lambda_j}(\lambda_j) = 0$. If $j = k$, a similar computation yields

$$\begin{aligned}
\langle v_j, v_j\rangle &= \langle b_{\lambda_1}b_{\lambda_2}\cdots b_{\lambda_{j-1}}\widetilde{c}_{\lambda_j},\ b_{\lambda_1}b_{\lambda_2}\cdots b_{\lambda_{j-1}}\widetilde{c}_{\lambda_j}\rangle \\
&= \langle \widetilde{c}_{\lambda_j}, \widetilde{c}_{\lambda_j}\rangle = 1.
\end{aligned}$$

Thus, $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal set in $\mathscr{K}_B$. Since $\dim \mathscr{K}_B = n$ (see (12.1.4)), we conclude that $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal basis for $\mathscr{K}_B$. □

12.3 Reproducing Kernel

Let $\mathscr{H}$ be a Hilbert space of analytic functions on a domain $\Omega \subseteq \mathbb{C}$. Then $\mathscr{H}$ is a *reproducing kernel Hilbert space* if there is a *kernel*

$$K(z, \lambda) : \Omega \times \Omega \to \mathbb{C}$$

such that for each fixed $\lambda \in \Omega$, the function

$$k_\lambda(z) = K(z, \lambda) \tag{12.3.1}$$

belongs to $\mathscr{H}$ and has the *reproducing property*

$$f(\lambda) = \big\langle f(\cdot), K(\cdot, \lambda)\big\rangle_{\mathscr{H}}, \qquad \lambda \in \Omega, \ f \in \mathscr{H}.$$

For example, the Hardy space H^2 (see Appendix A.4) is a reproducing kernel Hilbert space with kernel

$$K(z, \lambda) = \frac{1}{1 - \overline{\lambda} z}. \tag{12.3.2}$$

In particular, the kernel for the Hardy space is the Cauchy kernel encountered in (12.1.11). More information about reproducing kernel Hilbert spaces can be found in the texts [1, 111].

Let B be a finite Blaschke product and let

$$k_\lambda(z) := \frac{1 - \overline{B(\lambda)} B(z)}{1 - \overline{\lambda} z}. \tag{12.3.3}$$

The following result shows that $\mathscr{K}_B$ is a reproducing kernel Hilbert space with kernel $K(z, \lambda) = k_\lambda(z)$. Before proceeding, we should remark that although $f(\lambda) = \langle f, c_\lambda \rangle$ for all $f \in \mathscr{K}_B$ and $\lambda \in \mathbb{D}$, the function (12.3.2) does not serve as a reproducing kernel for $\mathscr{K}_B$ since c_λ does not, in general, belong to $\mathscr{K}_B$.

Proposition 12.3.4 *Let B be a finite Blaschke product, let $\lambda \in \mathbb{D}$, and define $k_\lambda(z)$ by* (12.3.3).

(a) $k_\lambda \in \mathscr{K}_B$.
(b) $\langle f, k_\lambda \rangle = f(\lambda)$ *for* $f \in \mathscr{K}_B$.
(c) $\|k_\lambda\|^2 = \dfrac{1 - |B(\lambda)|^2}{1 - |\lambda|^2}$ *for* $\lambda \in \mathbb{D}$.

Proof

(a) Since $\overline{B} B = 1$ on $\mathbb{T}$, for any polynomial q we have

$$\begin{aligned}
\langle Bq, k_\lambda \rangle &= \Big\langle Bq, \frac{1 - \overline{B(\lambda)} B(z)}{1 - \overline{\lambda} z} \Big\rangle \\
&= \langle Bq, c_\lambda \rangle - B(\lambda) \langle Bq, B c_\lambda \rangle \\
&= \langle Bq, c_\lambda \rangle - B(\lambda) \langle q, c_\lambda \rangle \\
&= B(\lambda) q(\lambda) - B(\lambda) q(\lambda) \\
&= 0.
\end{aligned}$$

Proposition 12.1.17 implies that $k_\lambda \in \mathcal{K}_B$.

(b) For any $f \in \mathcal{K}_B$,

$$\begin{aligned}\langle f, k_\lambda \rangle &= \left\langle f, \frac{1 - \overline{B(\lambda)}B(z)}{1 - \overline{\lambda}z} \right\rangle \\ &= \langle f, c_\lambda \rangle - B(\lambda)\langle f, Bc_\lambda \rangle \\ &= f(\lambda) - B(\lambda)\langle f, Bc_\lambda \rangle.\end{aligned}$$

By Proposition 12.1.17, the inner product in the second term equals

$$\begin{aligned}\left\langle f, \frac{B}{1 - \overline{\lambda}z} \right\rangle &= \left\langle f, B\sum_{n=0}^{\infty} \overline{\lambda}^n z^n \right\rangle \\ &= \sum_{n=0}^{\infty} \overline{\lambda}^n \langle f, z^n B \rangle \\ &= 0.\end{aligned}$$

Note that the series above converges since $|\lambda| < 1$ and

$$|\langle f, z^n B \rangle| \leqslant \|f\| \|z^n B\| = \|f\|$$

by the Cauchy–Schwarz inequality. Thus, $f(\lambda) = \langle f, k_\lambda \rangle$.

(c) It follows from (a) and (b) that for each $\lambda \in \mathbb{D}$,

$$\begin{aligned}\|k_\lambda\|^2 &= \langle k_\lambda, k_\lambda \rangle \\ &= k_\lambda(\lambda) \\ &= \frac{1 - |B(\lambda)|^2}{1 - |\lambda|^2}.\end{aligned}$$

This completes the proof. □

Since the Takenaka basis $\{v_1, v_2, \ldots v_n\}$ is an orthonormal basis for $\mathcal{K}_B$ (Proposition 12.2.2), the reproducing property of k_λ (Proposition 12.3.4) implies that

$$\begin{aligned}k_\lambda(z) &= \sum_{j=1}^{n} \langle k_\lambda, v_j \rangle v_j(z) \\ &= \sum_{j=1}^{n} \overline{v_j(\lambda)} v_j(z).\end{aligned} \tag{12.3.5}$$

In fact, the same argument holds for any orthonormal basis of $\mathcal{K}_B$. See Corollary 12.8.4 for other natural orthonormal bases of $\mathcal{K}_B$.

One can also consider k_ξ for $\xi \in \mathbb{T}$. From (12.2.3) we observe that the values $v_1(\xi), v_2(\xi), \ldots, v_n(\xi)$ are well defined. In light of (12.3.5), we define

$$k_\xi(z) := \sum_{j=1}^{n} \overline{v_j(\xi)} v_j(z). \tag{12.3.6}$$

Proposition 12.3.7 *Let B be a finite Blaschke product and let $\xi \in \mathbb{T}$.*

(a) $k_\xi \in \mathcal{K}_B$.
(b) $\|k_\xi\| = \sqrt{|B'(\xi)|}$.
(c) $f(\xi) = \langle f, k_\xi \rangle$ *for every* $f \in \mathcal{K}_B$.

Proof

(a) The identity in (12.3.6) says that k_ξ is a linear combination of the basis vectors $v_1, v_2, \ldots, v_n$. We conclude that $k_\xi \in \mathcal{K}_B$.

(b) Suppose that B has zeros $\lambda_1, \lambda_2, \ldots, \lambda_n$, repeated according to multiplicity. The identity in (12.3.6) shows that

$$\begin{aligned}
\|k_\xi\|^2 &= \langle k_\xi, k_\xi \rangle \\
&= \sum_{j=1}^{n} |v_j(\xi)|^2 \\
&= \sum_{j=1}^{n} \frac{1 - |\lambda_j|^2}{|1 - \overline{\lambda_j}\xi|^2} && \text{(by (12.2.3))} \\
&= |B'(\xi)| && \text{(by (3.4.8))}.
\end{aligned}$$

(c) Since each $f \in \mathcal{K}_B$ is analytic in a neighborhood of $\mathbb{D}^-$ and $k_\lambda \to k_\xi$ uniformly on $\mathbb{T}$ as $\lambda \to \xi$,

$$\begin{aligned}
f(\xi) &= \lim_{\lambda \to \xi} f(\lambda) \\
&= \lim_{\lambda \to \xi} \langle f, k_\lambda \rangle \\
&= \langle f, k_\xi \rangle.
\end{aligned}$$

Indeed, uniform convergence permits the interchange of limit and integral that is implicit in the preceding computation. □

12.4 Projections onto Model Spaces

Let $\mathscr{H}$ be a complex Hilbert space. An *orthogonal projection* P is a bounded linear transformation $P : \mathscr{H} \to \mathscr{H}$ that is self-adjoint ($P = P^*$) and idempotent ($P^2 = P$). Such an operator has closed range and fixes each element of its range. Moreover, the kernel of an orthogonal projection is orthogonal to its range (see A.6.7).

By considering the boundary function $f : \mathbb{T} \to \mathbb{C}$ of each $f \in \mathscr{K}_B$, we may regard $\mathscr{K}_B$ as a finite-dimensional subspace of L^2. This permits us to consider the orthogonal projection $P_B : L^2 \to L^2$ whose range is $\mathscr{K}_B$. This projection is intimately related to the kernels

$$c_\lambda = \frac{1}{1 - \overline{\lambda} z} \qquad \text{and} \qquad k_\lambda = \frac{1 - \overline{B(\lambda)} B(z)}{1 - \overline{\lambda} z}.$$

Proposition 12.4.1 *Let B be a finite Blaschke product.*

(a) $(P_B f)(\lambda) = \langle f, k_\lambda \rangle$ *for each* $f \in L^2$.
(b) $P_B c_\lambda = k_\lambda$ *for each* $\lambda \in \mathbb{D}$.

Proof

(a) Since $P_B f \in \mathscr{K}_B$, the reproducing property of k_λ implies that

$$\begin{aligned} (P_B f)(\lambda) &= \langle P_B f, k_\lambda \rangle \\ &= \langle f, P_B k_\lambda \rangle \\ &= \langle f, k_\lambda \rangle. \end{aligned} \tag{12.4.2}$$

The preceding two equalities follow from the self-adjointness of P_B and the fact that $k_\lambda \in \mathscr{K}_B$, respectively.

(b) Apply (12.4.2) with $f = c_\lambda$ and deduce that

$$\begin{aligned} (P_B c_\lambda)(z) &= \langle c_\lambda, k_z \rangle \\ &= \overline{\langle k_z, c_\lambda \rangle} \\ &= \overline{k_z(\lambda)} \\ &= k_\lambda(z). \end{aligned}$$

The final line follows from (12.3.3); see also Exercise 12.7. □

12.5 Conjugation

Let

$$B(z) = \gamma \prod_{j=1}^{n} \frac{z - \lambda_j}{1 - \overline{\lambda_j} z}, \tag{12.5.1}$$

in which $\gamma \in \mathbb{T}$. From (12.1.2), we see that each $f \in \mathcal{K}_B$ is of the form

$$f = \frac{P}{R},$$

where $P \in \mathcal{P}_{n-1}$ and

$$R(z) = (1 - \overline{\lambda_1} z)(1 - \overline{\lambda_2} z) \cdots (1 - \overline{\lambda_n} z). \tag{12.5.2}$$

Define a conjugate-linear map $C : \mathcal{K}_B \to \mathcal{K}_B$ by

$$C\left(\frac{P}{R}\right) = \frac{P^{\#}}{R}, \tag{12.5.3}$$

in which

$$P^{\#}(z) = z^{n-1}\overline{P(1/\overline{z})}.$$

That is, $P^{\#}$ is the polynomial of degree at most $n-1$ obtained from P by conjugating its coefficients and reversing their order; see Sect. 3.3 for a review of the # operation. It follows from (12.1.2) that $Cf \in \mathcal{K}_B$ whenever $f \in \mathcal{K}_B$.

The map C is a *conjugation* on $\mathcal{K}_B$. That is, it is conjugate-linear, involutive, and isometric.

Proposition 12.5.4 *The map $C : \mathcal{K}_B \to \mathcal{K}_B$ has the following properties.*

(a) $C(\alpha f + g) = \overline{\alpha} Cf + Cg$ *for all* $f, g, \in \mathcal{K}_B$ *and* $\alpha \in \mathbb{C}$.
(b) $C^2 = I$.
(c) $|Cf| = |f|$ *on* $\mathbb{T}$ *for all* $f \in \mathcal{K}_B$.
(d) $\|Cf\| = \|f\|$ *for all* $f \in \mathcal{K}_B$.
(e) $\langle f, g\rangle = \langle Cg, Cf\rangle$ *for all* $f, g \in \mathcal{K}_B$.

Proof Let B and R be defined as in (12.5.1) and (12.5.2), respectively.

(a) This follows from the fact that the map $P \mapsto P^{\#}$ on $\mathcal{P}_{n-1}$ is conjugate linear.
(b) Since $(P^{\#})^{\#} = P$ by (3.3.7), it follows that $C^2(P/R) = C(P^{\#}/R) = P/R$ for all $P \in \mathcal{P}_{n-1}$. Thus, $C^2 = I$.
(c) If $f \in \mathcal{K}_B$, write $f = P/R$, in which $P \in \mathcal{P}_{n-1}$. Since $|P| = |P^{\#}|$ on $\mathbb{T}$ by (3.3.12), it follows that $|f| = |P/R| = |P^{\#}/R| = |Cf|$ on $\mathbb{T}$.

(d) From the preceding we see that

$$\begin{aligned}\|Cf\|^2 &= \int_0^{2\pi} |Cf(e^{i\theta})|^2 \frac{d\theta}{2\pi}\\ &= \int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi}\\ &= \|f\|^2.\end{aligned}$$

(e) See Exercise 12.6. □

For some purposes, it is more convenient to work on the unit circle than on the unit disk. This is particularly true for model spaces.

Proposition 12.5.5 *Let B be a finite Blaschke product.*

(a) If $f \in \mathcal{K}_B$, then

$$(Cf)(\zeta) = \overline{f(\zeta)\zeta} B(\zeta), \qquad \zeta \in \mathbb{T}. \tag{12.5.6}$$

(b) A rational function f that is analytic on a neighborhood of $\mathbb{D}^-$ belongs to $\mathcal{K}_B$ if and only if there is a rational function g that is analytic on a neighborhood of $\mathbb{D}^-$ so that $f(\zeta) = \overline{g(\zeta)\zeta} B(\zeta)$ on $\mathbb{T}$. If this occurs, then $Cf = g$.

Proof

(a) Let B denote the finite Blaschke product (12.5.1) and let R denote the polynomial from (12.5.2). Since B is of degree n, each $f \in \mathcal{K}_B$ can be written as $f = P/R$ for some $P \in \mathscr{P}_{n-1}$. By (3.3.9), we have $B = R^\#/R$, in which

$$R^\#(z) = (z - a_1)(z - a_2)\cdots(z - a_n).$$

Since $C(P/R) = P^\#/R$, where $P^\#(z) = z^{n-1}\overline{P(1/\overline{z})}$, it follows that

$$\begin{aligned}(Cf)(\zeta) &= \frac{P^\#(\zeta)}{R(\zeta)}\\ &= \frac{P^\#(\zeta)}{R^\#(\zeta)} \cdot \frac{R^\#(\zeta)}{R(\zeta)}\\ &= \frac{\zeta^{n-1}\overline{P(1/\overline{\zeta})}}{\zeta^n \overline{R(1/\overline{\zeta})}} \cdot B(\zeta)\\ &= \frac{\overline{P(\zeta)}}{\zeta \overline{R(\zeta)}} \cdot B(\zeta)\end{aligned}$$

$$= \overline{\left(\frac{P(\zeta)}{R(\zeta)}\zeta\right)} B(\zeta)$$
$$= \overline{f(\zeta)\zeta} B(\zeta)$$

for all $\zeta \in \mathbb{T}$. This establishes (12.5.6).

(b) If $f \in \mathcal{K}_B$, then (a) ensures that $g = Cf$ satisfies the desired condition. Conversely, suppose that f is a rational function that is analytic on a neighborhood of $\mathbb{D}^-$ and that g is another such rational function so that $f(\zeta) = \overline{g(\zeta)\zeta} B(\zeta)$ on $\mathbb{T}$. For all $k \geqslant 0$,

$$\begin{aligned}
\langle Bz^k, f\rangle &= \int_0^{2\pi} B(e^{i\theta})e^{ik\theta}\overline{f(e^{i\theta})}\,\frac{d\theta}{2\pi} \\
&= \int_0^{2\pi} B(e^{i\theta})e^{ik\theta}\overline{\overline{g(e^{i\theta})e^{i\theta}}B(e^{i\theta})}\,\frac{d\theta}{2\pi} \\
&= \int_0^{2\pi} e^{ik\theta}g(e^{i\theta})ie^{i\theta}\,\frac{d\theta}{2\pi i} \\
&= \frac{1}{2\pi i}\oint_{\mathbb{T}} \zeta^k g(\zeta)\,d\zeta \\
&= 0
\end{aligned}$$

by Cauchy's theorem. Proposition 12.1.17 implies that $f \in \mathcal{K}_B$. The computation in the proof of (a) confirms that $Cf = g$. □

Example 12.5.7 What is the conjugate of the kernel $k_\lambda \in \mathcal{K}_B$ defined in (12.3.3)? Since $B\overline{B} = 1$ on $\mathbb{T}$, for $\zeta \in \mathbb{T}$ we appeal to (12.5.6) and obtain

$$\begin{aligned}
(Ck_\lambda)(\zeta) &= \overline{\left(\frac{1-\overline{B(\lambda)}B(\zeta)}{1-\overline{\lambda}\zeta}\right)\overline{\zeta}} B(\zeta) \\
&= \frac{1-B(\lambda)\overline{B(\zeta)}}{1-\lambda\overline{\zeta}}\cdot\frac{B(\zeta)}{\zeta} \\
&= \frac{B(\zeta)-B(\lambda)}{\zeta-\lambda}.
\end{aligned}$$

The identity principle implies that

$$(Ck_\lambda)(z) = \frac{B(z)-B(\lambda)}{z-\lambda}$$

for all $z, \lambda \in \mathbb{D}$. In particular, for $B(0) = 0$ it follows that $k_0 \equiv 1$ and

$$(C1)(z) = \frac{B(z)}{z}. \tag{12.5.8}$$

The conjugation C acts on the Takenaka basis $\{v_1, v_2, \ldots, v_n\}$ as follows.

Proposition 12.5.9 $Cv_k = \begin{cases} (b_{\lambda_{\ell+1}} b_{\lambda_{\ell+2}} \cdots b_{\lambda_n})\widetilde{c}_{\lambda_\ell} & \textit{if } k = 1, 2, \ldots, n-1, \\ \widetilde{c}_{\lambda_n} & \textit{if } k = n. \end{cases}$

Proof In what follows, we compute on $\mathbb{T}$ with ζ as the independent variable. For $\ell = 1, 2, \ldots, n-1$, (12.2.3) and (12.5.6) yield

$$\begin{aligned} Cv_\ell &= \overline{v_\ell \zeta} B \\ &= \overline{(b_{\lambda_1} b_{\lambda_2} \cdots b_{\lambda_{\ell-1}} \widetilde{c}_{\lambda_\ell})\zeta} (b_{\lambda_1} b_{\lambda_2} \cdots b_{\lambda_n}) \\ &= b_{\lambda_\ell} b_{\lambda_{\ell+1}} \cdots b_{\lambda_n} \frac{\overline{\zeta}\sqrt{1 - |\lambda_\ell|^2}}{1 - \lambda_\ell \overline{\zeta}} \\ &= \sqrt{1 - |\lambda_\ell|^2}\, \frac{b_{\lambda_\ell} b_{\lambda_{\ell+1}} \cdots b_{\lambda_n}}{\zeta - \lambda_\ell} \\ &= \sqrt{1 - |\lambda_\ell|^2}\, b_{\lambda_{\ell+1}} b_{\lambda_{\ell+2}} \cdots b_{\lambda_n} c_{\lambda_\ell} \\ &= (b_{\lambda_{\ell+1}} b_{\lambda_{\ell+2}} \cdots b_{\lambda_n}) \widetilde{c}_{\lambda_\ell}. \end{aligned}$$

The proof that $Cv_n = \widetilde{c}_{\lambda_n}$ is similar; see Exercise 12.9. □

12.6 Compressed Shift

For the remainder of the chapter, we assume that B is a finite Blaschke product with

$$B(0) = 0. \tag{12.6.1}$$

This assumption simplifies many of the following computations. A consequence of our assumption follows from Proposition 12.2.2, which tells us that $v_1 \equiv 1 \in \mathscr{K}_B$.

For an analytic function f on $\mathbb{D}$, we define

$$(Sf)(z) = zf(z). \tag{12.6.2}$$

This is the *unilateral shift* operator (initially studied on the Hardy space H^2; see Appendix A.4). It is one of the most important objects in operator theory [59, 106]. We leave it to the reader to verify that the adjoint S^* of S on H^2 is given by

$$(S^* f)(z) = \frac{f(z) - f(0)}{z}.$$

Definition 12.6.3 The operator

$$S_B : \mathcal{K}_B \to \mathcal{K}_B, \quad S_B = P_B S|_{\mathcal{K}_B} \tag{12.6.4}$$

is the *compressed shift*.

The importance of this operator stems from the fact that it can be used to represent certain types of contractions. Before getting into the details, we first establish some basic properties of S_B. From Corollary 12.1.20 we know that

$$f \in \mathcal{K}_B \implies \frac{f - f(0)}{z} \in \mathcal{K}_B.$$

The following proposition asserts that S_B^* is the restriction of the backward shift operator (12.1.19) to $\mathcal{K}_B$; that is $S_B^* = S^*|_{\mathcal{K}_B}$.

Proposition 12.6.5 *For a finite Blaschke product B,*

$$S_B^* f = \frac{f - f(0)}{z}, \qquad f \in \mathcal{K}_B.$$

Proof For any $f, g \in \mathcal{K}_B$,

$$\begin{aligned}
\langle S_B^* f, g\rangle &= \langle f, S_B g\rangle \\
&= \langle f, P_B(zg)\rangle \\
&= \langle P_B f, zg\rangle \\
&= \langle f, zg\rangle \\
&= \langle f, zg\rangle - f(0)\langle 1, zg\rangle \qquad \text{(by Lemma A.2.2)} \\
&= \langle f - f(0), zg\rangle \\
&= \langle \overline{z}(f - f(0)), g\rangle \\
&= \Big\langle \frac{f - f(0)}{z}, g\Big\rangle.
\end{aligned}$$

Since this holds for all $f, g \in \mathcal{K}_B$, we obtain the desired identity. □

We can use the Takenaka basis $\{v_1, v_2, \ldots, v_n\}$ for $\mathcal{K}_B$, along with Proposition 12.4.1 and (A.6.8), to write the compressed shift S_B as

$$S_B f = \sum_{j=1}^{n} \langle zf, v_j\rangle v_j.$$

The matrix representation $[S_B]$ of S_B with respect to the Takenaka basis is then

$$[S_B] = [\langle z v_k, v_j\rangle]_{1 \leqslant j,k \leqslant n}.$$

Proposition 12.6.6 *If B is a finite Blaschke with zeros $\lambda_1, \lambda_2, \ldots, \lambda_n$, repeated according to multiplicity, then $[S_B]$ is lower triangular and has $\lambda_1, \lambda_2, \ldots, \lambda_n$ along its main diagonal.*

Proof For $j < k$,

$$\begin{aligned}
[S_B]_{j,k} &= \langle z v_k, v_j\rangle \\
&= \langle z \cdot b_{\lambda_1} b_{\lambda_2} \cdots b_{\lambda_{k-1}} \widetilde{c_{\lambda_k}},\ b_{\lambda_1} b_{\lambda_2} \cdots b_{\lambda_{j-1}} \widetilde{c_{\lambda_j}}\rangle \\
&= \sqrt{1 - |\lambda_j|^2}\sqrt{1 - |\lambda_k|^2}\langle z b_{\lambda_j} b_{\lambda_{j+1}} \cdots b_{\lambda_{k-1}} c_{\lambda_k},\ c_{\lambda_j}\rangle \\
&= \sqrt{1 - |\lambda_j|^2}\sqrt{1 - |\lambda_k|^2}\lambda_j \left(\prod_{i=j}^{k-1} \frac{\lambda_j - \lambda_i}{1 - \overline{\lambda_i}\lambda_j}\right) c_{\lambda_k}(\lambda_j) \\
&= 0.
\end{aligned}$$

Thus, $[S_B]$ lower triangular. The diagonal entries are

$$\begin{aligned}
[S_B]_{j,j} &= \langle z b_{\lambda_1} b_{\lambda_2} \cdots b_{\lambda_{j-1}} \widetilde{c_{\lambda_j}},\ b_{\lambda_1} b_{\lambda_2} \cdots b_{\lambda_{j-1}} \widetilde{c_{\lambda_j}}\rangle \\
&= (1 - |\lambda_j|^2)\langle z c_{\lambda_j}, c_{\lambda_j}\rangle \\
&= (1 - |\lambda_j|^2)\lambda_j c_{\lambda_j}(\lambda_j) \\
&= (1 - |\lambda_j|^2)\lambda_j \frac{1}{1 - |\lambda_j|^2} \\
&= \lambda_j.
\end{aligned}$$

This completes the proof. □

See Exercise 12.12 for more about the matrix representation of S_B.

Corollary 12.6.7 *The eigenvalues of S_B are $\lambda_1, \lambda_2, \ldots, \lambda_n$.*

We end this section with a relationship between the compressed shift S_B and the conjugation C on $\mathscr{K}_B$. The following asserts that the compressed shift S_B is a *complex symmetric operator*.

Proposition 12.6.8 $C S_B C = S_B^*$.

Proof For $f, g \in \mathscr{K}_B$, use the formula

$$(Cf)(\zeta) = \overline{\zeta f(\zeta)} B(\zeta)$$

for $\zeta \in \mathbb{T}$ (Proposition 12.5.5) along with Proposition 12.5.4 to show that

$$\begin{aligned}\langle CS_B^*Cf, g\rangle &= \langle Cg, S_B^*Cf\rangle \\ &= \langle Cg, S^*Cf\rangle \\ &= \langle SCg, Cf\rangle \\ &= \langle \zeta B\overline{\zeta g}, B\overline{\zeta f}\rangle \\ &= \langle \zeta\overline{g}, \overline{f}\rangle \\ &= \langle \zeta f, g\rangle \\ &= \langle Sf, g\rangle \\ &= \langle Sf, P_B g\rangle \\ &= \langle P_B Sf, g\rangle \\ &= \langle S_B f, g\rangle.\end{aligned}$$

Since this holds for all $f, g \in \mathscr{K}_B$, the desired identity follows. □

12.7 Partial Isometries

Definition 12.7.1 A bounded linear operator on a Hilbert space $\mathscr{H}$ is a *partial isometry* if A is isometric on $(\ker A)^\perp$; that is,

$$\|A\mathbf{x}\| = \|\mathbf{x}\|, \quad \mathbf{x} \in (\ker A)^\perp.$$

Example 12.7.2 Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthonormal basis for $\mathbb{C}^n$ and let $1 < k \leqslant n$. If we regard $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ as column vectors, then $A \in M_n$ defined by

$$A = [\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_k\ \mathbf{0}\ \mathbf{0}\ \ldots\ \mathbf{0}]$$

is a partial isometry. To be more precise, A is the matrix representation, with respect to the standard basis of $\mathbb{C}^n$, of a partial isometry. As a slight abuse of language, we will say that A is a partial isometry. Indeed,

$$\ker A = \operatorname{span}\{\mathbf{e}_{k+1}, \mathbf{e}_{k+2}, \ldots, \mathbf{e}_n\}$$

and

$$(\ker A)^\perp = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k\},$$

in which $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ denotes the standard basis for $\mathbb{C}^n$. If

$$\mathbf{z} = (z_1, z_2, \ldots, z_k, 0, 0, \ldots 0) \in (\ker A)^\perp,$$

then

$$A\mathbf{z} = z_1\mathbf{u}_1 + z_2\mathbf{u}_2 + \cdots + z_k\mathbf{u}_k.$$

The fact that the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ are orthonormal yields

$$\|A\mathbf{z}\|^2_{\mathbb{C}^n} = |z_1|^2 + |z_2|^2 + \cdots + |z_k|^2 = \|\mathbf{z}\|^2_{\mathbb{C}^n}.$$

The following proposition asserts that Example 12.7.2 is typical [71, Cor. 2]; see also Exercise 12.14. Recall the definition of unitary equivalence of operators from (A.7.7).

Proposition 12.7.3 *A partial isometry of rank k on a Hilbert space of degree n is unitarily equivalent to an $n \times n$ matrix of the form*

$$[\mathbf{u}_1 \ \mathbf{u}_2 \ \ldots \ \mathbf{u}_k \ \mathbf{0} \ \mathbf{0} \ \ldots \ \mathbf{0}]$$

for some (possibly empty) list of orthonormal column vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \in \mathbb{C}^n$.

We leave it to the reader to verify the following facts about partial isometries; see Exercise 12.13.

Proposition 12.7.4 *For $A \in \mathscr{L}(\mathscr{H})$ the following are equivalent.*

(a) A is a partial isometry.
(b) A^* is a partial isometry.
(c) $A = AA^*A$.
(d) A^*A is an orthogonal projection.
(e) AA^* is an orthogonal projection.

If A is a partial isometry, then $A^*A \in \mathscr{L}(\mathscr{H})$ is the orthogonal projection with range $(\ker A)^\perp$ and $AA^* \in \mathscr{L}(\mathscr{H})$ is the orthogonal projection with range $(\ker A^*)^\perp$.

Recall the linear transformation $\mathbf{x} \otimes \mathbf{y} : \mathscr{H} \to \mathscr{H}$ defined by

$$(\mathbf{x} \otimes \mathbf{y})(\mathbf{z}) := \langle \mathbf{z}, \mathbf{y} \rangle \mathbf{x}, \quad \mathbf{z} \in \mathscr{H}.$$

It has rank one if $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$. Its adjoint is $\mathbf{y} \otimes \mathbf{x}$ (see (A.6.9)).

In what follows, recall the hypothesis $B(0) = 0$ from (12.6.1). This ensures that B/z is analytic on $\mathbb{D}$. It belongs to $\mathscr{K}_B$ by Proposition 12.1.17.

Proposition 12.7.5 *If* $B(0) = 0$, *then*

$$S_B S_B^* = I - 1 \otimes 1 \qquad \textit{and} \qquad S_B^* S_B = I - \frac{B}{z} \otimes \frac{B}{z}. \tag{12.7.6}$$

Proof For each $f \in \mathscr{K}_B$,

$$\begin{aligned}
(I - S_B S_B^*) f &= f - P_B(SS^* f) && \text{(by Proposition 12.6.5)}\\
&= f - P_B(f - f(0))\\
&= f(0) P_B 1\\
&= \langle f, 1\rangle 1 && \text{(by Proposition 12.4.1)}\\
&= (1 \otimes 1) f.
\end{aligned}$$

This proves the first identity in (12.7.6). To prove the second, first show that

$$C(f \otimes g)C = Cf \otimes Cg;$$

see Exercise (12.10). Then

$$\begin{aligned}
C(1 - S_B S_B^*)C &= CC - C S_B S_B^* C\\
&= I - C S_B C C S_B^* C\\
&= I - S_B^* S_B && \text{(by Proposition 12.6.8)}
\end{aligned}$$

and hence

$$\begin{aligned}
I - S_B^* S_B &= C(I - S_B S_B^*)C\\
&= C(1 \otimes 1)C\\
&= (C1) \otimes (C1)\\
&= \frac{B}{z} \otimes \frac{B}{z} && \text{(by (12.5.8))}.
\end{aligned}$$

This completes the proof. □

Corollary 12.7.7 *If* $B(0) = 0$, *then* S_B *is a partial isometry with*

$$\ker S_B = \operatorname{span}\left\{\frac{B}{z}\right\} \qquad \textit{and} \qquad \ker S_B^* = \operatorname{span}\{1\}.$$

Proof Since 1 and B/z are unit vectors, $1 \otimes 1$ and $B/z \otimes B/z$ are orthogonal projections on $\mathscr{K}_B$ which ranges are $\operatorname{span}\{1\}$ and $\operatorname{span}\{B/z\}$ respectively. The conclusions now follow from Propositions 12.7.4 and 12.7.5. □

For a finite Blaschke product B with $B(0) = 0$, the compressed shift S_B is a partial isometry on a finite-dimensional space. Moreover, its kernel is one dimensional and its eigenvalues are contained in $\mathbb{D}$. It turns out that S_B is a "model" for such operators.

Theorem 12.7.8 *Suppose that*

(a) $A \in M_n$ *is a partial isometry;*
(b) $\dim \ker A = 1$*;*
(c) the eigenvalues of A *lie inside* $\mathbb{D}$*.*

If $\{0, \lambda_2, \lambda_3, \ldots, \lambda_n\}$ *are the eigenvalues of* A*, repeated according to multiplicity, and*

$$B(z) = z \prod_{j=2}^{n} \frac{\lambda_j - z}{1 - \overline{\lambda_j} z},$$

then A *is unitarily equivalent to* S_B*.*

This representation theorem is a consequence of the following result.

Theorem 12.7.9 (Halmos–McLaughlin [67]) *Suppose* U *and* V *are two partial isometries with one-dimensional kernels. Then* U *and* V *are unitarily equivalent if and only if they have the same eigenvalues and the same multiplicities.*

Proof We choose to work with matrices instead of operators here. If $U, V \in M_n$ are unitarily equivalent, then they are similar and hence have the same characteristic polynomials. Thus, U and V have the same eigenvalues with the same multiplicities.

Now for the converse. We proceed by induction on n. If $n = 1$, then each 1×1 partial isometry with one-dimensional kernel is the 1×1 zero matrix $[0]$. Suppose for our induction hypothesis that each pair of $n \times n$ partial isometries sharing the same eigenvalues and multiplicities is unitarily equivalent.

Let U and V be two $(n+1) \times (n+1)$ partial isometries with the same eigenvalues and multiplicities. Schur's theorem on unitary triangularization (Theorem A.8.1) permits us, via unitary equivalence, to assume that U and V are upper-triangular matrices of the form

$$U = \begin{bmatrix} U' & \mathbf{u} \\ \mathbf{0}^* & \alpha \end{bmatrix} \qquad \text{and} \qquad V = \begin{bmatrix} V' & \mathbf{v} \\ \mathbf{0}^* & \alpha \end{bmatrix}, \tag{12.7.10}$$

in which $\alpha \in \mathbb{C}$, $\mathbf{u}, \mathbf{v}, \mathbf{0} \in \mathbb{C}^n$, and U', V' are $n \times n$ upper-triangular matrices with $0, \lambda_2, \lambda_3, \ldots, \lambda_n$ (in that order) along their main diagonals.

Since U is upper triangular and has a 0 in the $(1, 1)$ position, we may write

$$U = [\mathbf{0}\ \mathbf{c}_2\ \mathbf{c}_3\ \ldots\ \mathbf{c}_{n+1}] \tag{12.7.11}$$

in column-by-column format. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n+1}$ denote the standard basis vectors for $\mathbb{C}^{n+1}$. By hypothesis, $\dim \ker U = 1$ and hence $\ker U = \operatorname{span}\{\mathbf{e}_1\}$. Since U is a partial isometry, it is isometric on

$$(\ker U)^{\perp} = \operatorname{span}\{\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_{n+1}\} \tag{12.7.12}$$

and hence

$$\Big\| U\Big(\sum_{j=2}^{n+1} z_j \mathbf{e}_j\Big)\Big\|^2 = \Big\|\sum_{j=2}^{n+1} z_j \mathbf{e}_j\Big\|^2 = \sum_{j=2}^{n+1} |z_j|^2.$$

for any $z_2, z_3, \ldots, z_{n+1} \in \mathbb{C}$. We also have

$$U\Big(\sum_{j=2}^{n+1} z_j \mathbf{e}_j\Big) = [\mathbf{0}\ \mathbf{c}_2\ \mathbf{c}_3\ \ldots\ \mathbf{c}_{n+1}] \begin{bmatrix} 0 \\ z_2 \\ \vdots \\ z_{n+1} \end{bmatrix} = \sum_{j=2}^{n+1} z_i \mathbf{c}_j.$$

Using the previous equation, along with (12.7.12), it follows that

$$\sum_{j=2}^{n+1} |z_j|^2 = \Big\| U\Big(\sum_{j=2}^{n+1} z_j \mathbf{e}_j\Big)\Big\|^2 = \sum_{2 \leqslant j,k \leqslant n+1} z_j \overline{z_k} \langle \mathbf{c}_j, \mathbf{c}_k \rangle \tag{12.7.13}$$

for all $z_2, z_3, \ldots, z_{n+1} \in \mathbb{C}$. Thus, $\{\mathbf{c}_2, \mathbf{c}_3, \ldots, \mathbf{c}_{n+1}\}$ is an orthonormal basis for $(\ker U)^{\perp}$. In particular, since U is upper triangular, the matrix U' from (12.7.10) takes the form

$$U' = [\mathbf{0}\ \mathbf{q}_2\ \ldots\ \mathbf{q}_{n+1}], \tag{12.7.14}$$

in which $\{\mathbf{q}_2, \mathbf{q}_3, \ldots, \mathbf{q}_{n+1}\}$ is an orthonormal set and $\mathbf{u}$ is orthogonal to the columns of U'. Proposition 12.7.3 ensures that U' is a partial isometry. From (12.7.14), we conclude that U' has a one-dimensional kernel. An analogous argument shows that V' has the same properties.

By our induction hypothesis, U' and V' are unitarily equivalent, and so

$$W_0 U_0 W_0^* = V_0$$

for some $n \times n$ unitary matrix W_0. Fix $\xi \in \mathbb{T}$ and consider the $(n+1) \times (n+1)$ matrix

$$W_\xi = \begin{bmatrix} W_0 & \mathbf{0} \\ \mathbf{0}^* & \xi \end{bmatrix}.$$

Then

$$W_\xi U W_\xi^* = \begin{bmatrix} V' & \overline{\xi} W_0 \mathbf{u} \\ \mathbf{0}^* & \alpha \end{bmatrix}$$

is a partial isometry of rank n whose first column is zero. Consequently, $\overline{\xi} W_0 \mathbf{u}$ is orthogonal to the columns of V'. Since $\mathbf{v}$ is also orthogonal to the columns of V', it follows that

$$\overline{\xi} W_0 \mathbf{u} = c\mathbf{v} \tag{12.7.15}$$

for some $c \in \mathbb{C}$. From (12.7.10),

$$\begin{bmatrix} \mathbf{u} \\ \alpha \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{v} \\ \alpha \end{bmatrix}$$

are unit vectors and hence $\|\mathbf{u}\| = \|\mathbf{v}\|$. Since W_0 is unitary and $|\xi| = 1$, we conclude from (12.7.15) that

$$\|\mathbf{u}\| = \|W'\mathbf{u}\| = |c|\|\mathbf{v}\|,$$

so $|c| = 1$. Hence there is a $\xi \in \mathbb{T}$ such that $\overline{\xi} W_0 \mathbf{u} = \mathbf{v}$. With this ξ we have

$$W_\xi U W_\xi^* = \begin{bmatrix} V' & \mathbf{v} \\ \mathbf{0}^* & \alpha \end{bmatrix} = V,$$

which says that U is unitarily equivalent to V. This completes the induction. □

12.8 Unitary Extensions of the Compressed Shift

A partial isometry on a finite-dimensional Hilbert space $\mathscr{H}$ can always be extended to a unitary operator on $\mathscr{H}$ [71]; see Exercise 12.15. That is, there is a unitary operator $U : \mathscr{H} \to \mathscr{H}$ so that

$$U|_{(\ker V)^\perp} = V.$$

This is no longer the case if one considers partial isometries on infinite-dimensional spaces: the unilateral shift (12.6.2) is a partial isometry that has no unitary extensions.

For a finite Blaschke product B with $B(0) = 0$, we observed in Corollary 12.7.7 that S_B is a partial isometry on the model space $\mathscr{K}_B$. In this section we identify all unitary extensions of S_B and discuss their eigenvalues and eigenvectors. This work of D. Clark [24] is valid in a much broader setting than we present here; see [59] for more details.

For a fixed $\alpha \in \mathbb{T}$ define $U_\alpha : \mathscr{K}_B \to \mathscr{K}_B$ by

$$U_\alpha := S_B + \alpha\Big(1 \otimes \frac{B}{z}\Big). \tag{12.8.1}$$

In other words,

$$U_\alpha f = S_B f + \alpha\Big\langle f, \frac{B}{z}\Big\rangle, \qquad f \in \mathscr{K}_B.$$

Theorem 12.8.2 *If B is a finite Blaschke product with $B(0) = 0$ and $\alpha \in \mathbb{T}$, then U_α is unitary on $\mathscr{K}_B$.*

Proof Since $\mathscr{K}_B$ is finite dimensional, a left inverse of U_α is also a right inverse of U_α. Thus, it suffices to show that

$$U_\alpha^* U_\alpha = I.$$

We use (A.6.10) and compute

$$\begin{aligned}
U_\alpha^* U_\alpha &= \bigg(S_B + \alpha\Big(1 \otimes \frac{B}{z}\Big)\bigg)^* \bigg(S_B + \alpha\Big(1 \otimes \frac{B}{z}\Big)\bigg)\\
&= \bigg(S_B^* + \overline{\alpha}\Big(1 \otimes \frac{B}{z}\Big)^*\bigg)\bigg(S_B + \alpha\Big(1 \otimes \frac{B}{z}\Big)\bigg)\\
&= \bigg(S_B^* + \overline{\alpha}\Big(\frac{B}{z} \otimes 1\Big)\bigg)\bigg(S_B + \alpha\Big(1 \otimes \frac{B}{z}\Big)\bigg)\\
&= S_B^* S_B + \overline{\alpha}\Big(\frac{B}{z} \otimes 1\Big)S_B + \alpha S_B^*\Big(1 \otimes \frac{B}{z}\Big) + \Big(\frac{B}{z} \otimes 1\Big)\Big(1 \otimes \frac{B}{z}\Big).
\end{aligned}$$

By Corollary 12.7.7, the range of S_B is orthogonal to the span of the constant function 1. Thus,

$$\Big(\frac{B}{z} \otimes 1\Big)S_B = 0.$$

Appealing to Corollary 12.7.7 again, we have $\ker S_B^* = \operatorname{span}\{1\}$, which yields

$$S_B^*\Big(1 \otimes \frac{B}{z}\Big) = 0.$$

Since $\langle 1, 1\rangle = 1$, a short computation shows that

$$\left(\frac{B}{z} \otimes 1\right)\left(1 \otimes \frac{B}{z}\right) = \frac{B}{z} \otimes \frac{B}{z}.$$

Finally observe from Proposition 12.7.5 that

$$S_B^* S_B = I - \frac{B}{z} \otimes \frac{B}{z},$$

which proves that $U_\alpha^* U_\alpha = I$. □

The operators U_α for $\alpha \in \mathbb{T}$ are the *Clark unitary operators*. Since each U_α is unitary, its eigenvalues are contained in $\mathbb{T}$. We now compute them, along with their corresponding eigenvectors. In what follows, we need the boundary reproducing kernels k_ξ for $\xi \in \mathbb{T}$, along with their basic properties; see Proposition 12.3.7. We follow the proof from [59, p. 237].

Theorem 12.8.3 *Let B be a finite Blaschke product of degree n. For each $\alpha \in \mathbb{T}$, the eigenvalues of U_α are the distinct solutions $\xi_1, \xi_2, \ldots, \xi_n$ to $B(\xi) = \alpha$. The corresponding eigenvectors are $k_{\xi_1}, k_{\xi_2}, \ldots, k_{\xi_n}$ and they form an orthogonal basis for $\mathcal{K}_B$.*

Proof Theorem 3.4.10 ensures that for each fixed $\alpha \in \mathbb{T}$, the equation

$$B(\xi) = \alpha$$

has n distinct solutions $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{T}$. Since U_α is unitary its eigenvalues are unimodular. Moreover, $U_\alpha f = \xi f$ if and only if $U_\alpha^* f = \overline{\xi} f$. The computation

$$\begin{aligned} U_\alpha^* f &= S_B^* f + \overline{\alpha}\left(\frac{B}{z} \otimes 1\right) f \\ &= \frac{f - f(0)}{z} + \overline{\alpha}\langle f, 1\rangle \frac{B}{z} \\ &= \frac{f - f(0)}{z} + \overline{\alpha} f(0) \frac{B}{z} \end{aligned}$$

implies that $U_\alpha^* f = \overline{\xi} f$ if and only if

$$\overline{\xi} f = \frac{f - f(0)}{z} + \overline{\alpha} f(0) \frac{B}{z}.$$

This happens precisely when

$$f = f(0) \frac{1 - \overline{\alpha} B}{1 - \overline{\xi} z}$$

$$= f(0)\frac{1 - \overline{B(\xi_j)}B}{1 - \overline{\xi_j}z}$$

$$= f(0)k_{\xi_j};$$

that is, f is a constant multiple of k_{ξ_j}. Because the eigenvalues $\xi_1, \xi_2, \ldots, \xi_n$ are distinct, the corresponding eigenvectors $k_{\xi_1}, k_{\xi_2}, \ldots, k_{\xi_n}$ are orthogonal. By (12.1.4), $\dim \mathcal{K}_B = n$, which says that these eigenvectors form an orthogonal basis. □

For a unitary operator, eigenvectors corresponding to distinct eigenvalues are orthogonal. Consequently, Proposition 12.3.7 provides the following corollary.

Corollary 12.8.4 *Let B be a finite Blaschke product of degree n and let $\alpha \in \mathbb{T}$. Denote by $\xi_1, \xi_2, \ldots, \xi_n$ the distinct solutions to $B(\xi) = \alpha$. Then the functions*

$$e_j = \frac{k_{\xi_j}}{\sqrt{|B'(\xi_j)|}}, \qquad 1 \leqslant j \leqslant n,$$

form an orthonormal basis for $\mathcal{K}_B$.

The basis $\{e_1, e_2, \ldots, e_n\}$ is called a *Clark basis* for $\mathcal{K}_B$.

12.9 Notes

Model Spaces

The characterization of $\mathcal{K}_B$ provided by Proposition 12.1.17 can be greatly generalized. If u is an inner function, then the corresponding model space $\mathcal{K}_u$ is the orthogonal complement of uH^2 in the Hardy space H^2. That is, $f \in H^2$ belongs to $\mathcal{K}_u$ if and only if f is orthogonal to uh for all $h \in H^2$. Since the polynomials are dense in H^2, this is equivalent to insisting that $\langle f, uz^k \rangle = 0$ for $k \geqslant 0$.

Partial Isometries

For more about partial isometries on finite dimensional spaces, see [41, 71]. There are also some results about operators that are similar to a partial isometry [57]. The analogue of Theorem 12.7.9 is not always true when the kernel of the partial isometry is not one dimensional. Indeed, the matrices

$$A = \begin{bmatrix} 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \end{bmatrix}$$

are partial isometries (verify that $A = AA^*A$ and $B = BB^*B$ and then use Proposition 12.7.4) whose characteristic polynomials are both equal to z^4. However, A and B are not unitarily equivalent (they are not even similar) since they have different Jordan canonical forms. Note that both A and B have two-dimensional kernels.

It turns out that for partial isometries with N-dimensional kernels, where $N \geqslant 1$, there is an $N \times N$ matrix-valued analytic function on $\mathbb{D}$, called the *Livšic characteristic function* [97], that determines when partial isometries (whose eigenvalues all lie in $\mathbb{D}$) are unitarily equivalent. When the $n \times n$ matrix has one-dimensional kernel with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (all contained in $\mathbb{D}$ and counted with multiplicity), then the Livšic characteristic function turns out to be

$$\prod_{j=1}^{n} \frac{z - \lambda_j}{1 - \overline{\lambda_j} z},$$

the finite Blaschke product whose zeros are the λ_j. For the matrices A and B above, the Livšic characteristic functions Λ_A and Λ_B (which will be 2×2 matrix-valued analytic functions on $\mathbb{D}$ since the kernels are two dimensional) are

$$\Lambda_A(z) = \begin{bmatrix} z & 0 \\ 0 & z^3 \end{bmatrix} \quad \text{and} \quad \Lambda_B = \begin{bmatrix} z^2 & 0 \\ 0 & z^2 \end{bmatrix}.$$

The functions Λ_A and Λ_B turn out to be “significantly different” (in a sense defined by Livšic) and reveal that the partial isometries A and B are not unitarily equivalent.

The Commutant

Theorem 12.7.8 shows that the compressed shift S_B on the model space $\mathscr{K}_B$ serves as a model for a certain class of partial isometries. One can identify the commutant

$$\{S_B\}' = \{T \in \mathscr{L}(\mathscr{K}_B) : TS_B = S_B T\}$$

of S_B via one the crowning achievements of operator theory, the *commutant lifting theorem* [59, p. 229].

Conjugations

The subject of conjugations goes far beyond model spaces and appears in many contexts (with many applications) in operator theory [54, 55, 58].

Model Spaces

Model spaces are important in operator theory and complex analysis. For one, they have infinite-dimensional analogues and, like in our Halmos–McLaughlin presentation, the associated compressed shift is a model operator for certain types of contractions. Good sources are [59, 106, 131]. In particular, Theorem 12.7.8 can be expanded to the following. Suppose $A \in M_n$ is a contraction, the eigenvalues of A are contained in $\mathbb{D}$, and $I - A^*A$ has rank one. Then A is unitarily equivalent to the compressed shift S_B for some finite Blaschke product B.

Numerical Range

The numerical range $W(S_B)$ of the compressed shift S_B has been discussed in [21, 29, 62]. In particular, there is the following result. If B is a finite Blaschke product, let $B_1 = zB$. For each $\theta \in [0, 2\pi]$, let F_θ denote the convex hull of the solutions to $B_1(z) = e^{i\theta}$. Then

$$W(S_B) = \bigcap_{\theta \in [0,2\pi]} F_\theta.$$

Observe that the eigenvalues of S_B are the zeros of B (Corollary 12.6.7). Furthermore, each F_θ contains the zeros of B_1 (Theorem 5.2.8), which, in turn, contains the zeros of B. This illustrates the fact that the numerical range of an operator contains the eigenvalues of the operator (Proposition 10.3.2). We should mention an upcoming book *Finding Ellipses: What Blaschke Products, Poncelet's Theorem and the Numerical Range Know about Each Other* by Gorkin, Daepp, Shaffer, and Voss that will cover in greater detail the beautiful geometry surrounding the numerical range of a compressed shift.

Clark Theory

There is a well-developed theory of D. Clark concerning unitary extensions of the compressed shift S_B beyond the finite Blaschke product case covered in this chapter.

There are many technicalities to overcome since the kernel functions k_ξ for $\xi \in \mathbb{T}$ are not always well defined when B is an infinite Blaschke product or, more generally, an inner function. Nevertheless, there is a lot one can say and the theory is beautiful and appears in a variety of settings. The original source is [24] while a more recent treatment can be found in [59].

12.10 Exercises

12.1 Prove that $\langle f, c_a^{(j)} \rangle = f^{(j)}(a)$ for all $f \in \mathscr{K}_B$.

12.2 Prove Proposition 12.1.16 in the general case, in which $m_1, m_2, \ldots, m_r$ are no longer assumed to all equal 1.

12.3 Prove Proposition 12.1.17 in the general case by using the Hermite interpolation theorem [87] and the Leibniz formula

$$\frac{d^j}{dz^j}(z^k B) = \sum_{s=0}^{j} \binom{j}{s} B^{(s)} \frac{d^{j-s}}{dz^{j-s}} z^k.$$

12.4 For two finite Blaschke products B_1 and B_2, show that $\mathscr{K}_{B_1} \subseteq \mathscr{K}_{B_2}$ if and only if B_1 divides B_2.

12.5 For the reader familiar with Toeplitz operators on the Hardy space H^2 (see Appendix A.7), show that for any finite Blaschke product B, the model space $\mathscr{K}_B$ is the kernel of the Toeplitz operator $T_{\overline{B}}$.

12.6 (a) Prove the polarization identity

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\big(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2 + i\|\mathbf{u}+i\mathbf{v}\|^2 - i\|\mathbf{u}-i\mathbf{v}\|^2\big),$$

valid for vectors $\mathbf{u}, \mathbf{v}$ in any complex Hilbert space. (b) Use the polarization identity to prove Proposition 12.5.4.d.

12.7 Let $K(z, \lambda)$ be a reproducing kernel on a complex Hilbert space $\mathscr{H}$. Prove that $K(z, \lambda) = \overline{K(\lambda, z)}$.

12.8 Let $\mathscr{H}$ be a complex Hilbert space. Show that if $\langle \mathbf{x}, \mathbf{h} \rangle = \langle \mathbf{y}, \mathbf{h} \rangle$ for all $\mathbf{h} \in \mathscr{H}$, then $\mathbf{x} = \mathbf{y}$. This principle is used in the proof of Proposition 12.4.1.

12.9 Complete the proof of Proposition 12.5.9 by showing that $Cv_n = \widetilde{c}_{\lambda_n}$.

12.10 Show that $C(f \otimes g)C = Cf \otimes Cg$ for all $f, g \in \mathscr{K}_B$.

12.11 Use Proposition 12.6.5 to prove that

$$S_B^n = P_B S^n|_{\mathscr{K}_B}, \quad n = 1, 2, \ldots.$$

12.12 Extend the statement of Proposition 12.6.6 to show that the matrix representation of S_B with respect to the Takenaka basis is

$$\begin{bmatrix} a_1 & & & & \\ & a_2 & & & \\ & & \ddots & & \\ & q_{j,k} & & a_{n-1} & \\ & & & & a_n \end{bmatrix},$$

where

$$q_{j,k} = \left(\prod_{i=j+1}^{k-1} (-\overline{a_i})\right) \sqrt{1 - |a_j|^2}\sqrt{1 - |a_k|^2}.$$

12.13 Verify Proposition 12.7.4.

12.14 Prove that for $V \in M_n$, the following are equivalent.

(a) V is a partial isometry.
(b) $V = Q[\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_r\ \mathbf{0}\ \mathbf{0}\ \ldots\ \mathbf{0}]Q^*$, where $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r : 1 \leqslant r \leqslant n\}$ is a (possibly empty) set of orthonormal vectors in $\mathbb{C}^n$ and Q is a unitary matrix.
(c) $V = UP$, where U is a unitary matrix and P is an orthogonal projection.

12.15 Use Proposition 12.7.3 to show that a partial isometry on a finite-dimensional Hilbert space has a unitary extension.

Chapter 13
The Darlington Synthesis Problem

The (scalar-valued) *Darlington synthesis problem* from electrical network theory asks the following question. Given $a \in H^\infty$, do there exist $b, c, d \in H^\infty$ such that the matrix-valued analytic function

$$U = \begin{bmatrix} a & -b \\ c & d \end{bmatrix} \tag{13.0.1}$$

is unitary almost everywhere on $\mathbb{T}$? That is, do there exist $b, c, d \in H^\infty$ so that

$$U(\zeta)^{-1} = U^*(\zeta),$$

for almost every $\zeta \in \mathbb{T}$? Recall that Fatou's theorem (Theorem A.3.1) ensures that functions in H^∞ have well-defined radial limits almost everywhere on $\mathbb{T}$. The negative sign in the $(1, 2)$ entry of the matrix in (13.0.1) is for notational convenience and is largely inconsequential.

In the early 1970s, Arov [3], and independently, Douglas and Helton [33], showed that a solution to the Darlington synthesis problem exists if and only if $\|a\|_\infty \leqslant 1$ and a is "pseudocontinuable of bounded type," a sort of generalized analytic continuation that originated in a seminal paper of Douglas, Shapiro, and Shields [34]. A survey of generalized analytic continuation can be found in [121].

In this chapter, we study the Darlington synthesis problem with *rational* data $a \in H^\infty$. In addition to providing an algorithm to construct all rational solutions, this work involves computations with finite Blaschke products and their corresponding model spaces. The model-space perspective also reveals an interesting quaternionic structure to the problem. The following material originates in [51, 52], both of which also address the more general case when a belongs to H^∞ but is no longer assumed to be rational.

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8_13

13.1 Factorization of Rational Functions

In what follows, we often let u denote a finite Blaschke product. Although we have used B heavily for this purpose in the past, it turns out that u is more appropriate for this subject matter since $u = \det U$ turns out to be a finite Blaschke product whenever U is an analytic matrix-valued rational function on $\mathbb{D}$ that is unitary-valued on $\mathbb{T}$ (see Sect. 13.3).

If f is a rational function that is analytic on $\mathbb{D}$, then it has a finite number of zeros there. If u is a finite Blaschke product with these zeros, repeated according to multiplicity, then

$$F = \frac{f}{u}$$

is a rational function with no zeros on $\mathbb{D}$ and

$$f = uF. \tag{13.1.1}$$

This factorization of a rational function is a special case of the Nevanlinna factorization of Smirnov functions [38].

Definition 13.1.2 If f is a rational function that is analytic on $\mathbb{D}$ and factored as in (13.1.1), then u is the *inner factor* of f and F is the *outer factor*. If u is a unimodular constant function, then f is a *rational outer function*.

Although we frequently use "the" when we refer to inner or outer factors, it is technically inappropriate according to our definition (Definition 3.1.2) of a finite Blaschke product. Indeed, we have said that any finite product of disk automorphisms is a finite Blaschke product. More advanced texts, which treat infinite Blaschke products, require certain normalizations before a function can be called a "Blaschke product" or an "outer function" [38, 106]. In light of Definition 13.1.2, it is more precise to say that u and F are determined up to offsetting unimodular constant factors since $uF = (\zeta u)(\overline{\zeta} F)$ for all $\zeta \in \mathbb{T}$. For convenience, however, we continue our relentless abuse of the article "the" in what follows.

Example 13.1.3 Let $\lambda_1 \in \mathbb{D}$ and $\lambda_2, \lambda_3 \in \mathbb{C}\backslash\mathbb{D}^-$, and define

$$f(z) = \frac{(z-\lambda_1)(z-\lambda_2)}{z-\lambda_3}.$$

Then

$$f(z) = \frac{z-\lambda_1}{1-\overline{\lambda_1}z}\left(\frac{(1-\overline{\lambda_1}z)(z-\lambda_2)}{z-\lambda_3}\right) = u(z)F(z),$$

where

$$F(z) = \frac{(1-\overline{\lambda_1}z)(z-\lambda_2)}{z-\lambda_3}.$$

Observe that the inner factor u is a degree-one Blaschke product and F is a rational function with no poles or zeros in $\mathbb{D}$; that is, a rational outer function.

The inner–outer factorization of a rational analytic function on $\mathbb{D}$ is unique, up to unimodular constant factors. Suppose that $u_1 F_1 = u_2 F_2$ are two such factorizations. Since F_1 and F_2 are rational functions that do not vanish on $\mathbb{D}$, the zeros of u_1 and u_2 are the same and they have the same multiplicities. Thus, $u_1 = \zeta u_2$ for some $\zeta \in \mathbb{T}$ and hence $F_1 = \overline{\zeta} F_2$.

The following important lemma tells us that the outer factor of a rational function that is analytic on $\mathbb{D}$ is determined by the modulus of the function on $\mathbb{T}$.

Lemma 13.1.4 *If f, g are rational functions that are analytic on $\mathbb{D}$ and $|f| = |g|$ on $\mathbb{T}$, then $f = B_f F$ and $g = B_g F$ for some finite Blaschke products B_f and B_g and some rational outer function F.*

Proof If either f or g has zeros or poles on $\mathbb{T}$, then the identity $|f| = |g|$ on $\mathbb{T}$ shows that these zeros or poles (on $\mathbb{T}$) must be of the same order. Consequently, f/g has a removable singularity at such points and hence f/g is a continuous, unimodular, rational function on $\mathbb{T}$. Corollary 3.5.4 produces two finite Blaschke products B_1 and B_2 so that $f/g = B_1/B_2$. The inner–outer factorization of $B_2 f = B_1 g$ provides a finite Blaschke product B and a rational outer function F so that

$$BF = B_2 f = B_1 g.$$

The zeros of $B_2 f$ and $B_1 g$ are among those of B, so $B_f = B/B_2$ and $B_g = B/B_1$ are finite Blaschke products. Thus, $f = B_f F$ and $g = B_g F$. □

Example 13.1.5 Let u denote a finite Blaschke product and let $k_\lambda \in \mathcal{K}_u$ denote the kernel function (12.3.3). By Proposition 12.5.4, $|Ck_\lambda| = |k_\lambda|$ on $\mathbb{T}$ (which can also be verified directly). According to the preceding lemma, k_λ and Ck_λ share a common outer factor. Since k_λ is rational, analytic on $\mathbb{D}$, and does not vanish on $\mathbb{D}$, it is outer. Thus, we expect that Ck_λ is a finite Blaschke product times k_λ. Let us verify this. As in (12.2.1), let

$$b_\lambda(z) = \frac{z-\lambda}{1-\overline{\lambda}z}.$$

Picking up where Example 12.5.7 left off, we deduce that

$$\begin{aligned}(Ck_\lambda)(z) &= \frac{u(\zeta)-u(\lambda)}{\zeta-\lambda}\\ &= \frac{u(\zeta)-u(\lambda)}{1-\overline{u(\lambda)}u(\zeta)}\cdot\frac{1-\overline{\lambda}\zeta}{\zeta-\lambda}\cdot\frac{1-\overline{u(\lambda)}u(\zeta)}{1-\overline{\lambda}\zeta}\\ &= \frac{(b_{u(\lambda)}\circ u)(\zeta))}{b_\lambda(\zeta)}k_\lambda(\zeta). \end{aligned} \tag{13.1.6}$$

Now observe that $b_{u(\lambda)}\circ u$ is a finite Blaschke product (Theorem 3.6.2) that vanishes at λ. Consequently, it is divisible by b_λ and hence the factor in front of $k_\lambda(z)$ in (13.1.6) is a finite Blaschke product. We conclude from this that k_λ and Ck_λ share the outer factor k_λ.

13.2 Finite Blaschke Products as Divisors in Model Spaces

Let u be a finite Blaschke product of degree n. Proposition 12.5.4 tells us that $|Cf| = |f|$ on $\mathbb{T}$ for all $f \in \mathcal{K}_u$. Thus, Lemma 13.1.4 yields finite Blaschke products B_f and B_{Cf}, along with a rational outer function F, so that

$$f = B_f F \qquad \text{and} \qquad Cf = B_{Cf}F. \tag{13.2.1}$$

Proposition 12.5.5 ensures that

$$Cf = \overline{fz}u$$

on $\mathbb{T}$. Thus,

$$B_{Cf}F = \overline{B_f F z}u$$

and hence

$$B_f B_{Cf} = \frac{\overline{Fz}u}{F}$$

on the subset of $\mathbb{T}$ where F is nonzero; that is, at all but finitely many points. Consequently, the finite Blaschke product

$$\mathrm{i}_F := B_f B_{Cf} \tag{13.2.2}$$

depends only upon F and u. It does not depend upon the particular pair of conjugate functions in $\mathcal{K}_u$ with common outer factor F that are chosen. If u is fixed and there is no chance of confusion, we say that i_F is the *finite Blaschke product associated with* F. On $\mathbb{T}$, it satisfies

$$\mathrm{i}_F F = \overline{Fz}u. \tag{13.2.3}$$

In the lattice of finite Blaschke products that occur as factors of functions in $\mathcal{K}_B$, the function i_F has the following maximality property.

Proposition 13.2.4 *Let u and B be finite Blaschke products and let F be the outer factor of a function in $\mathcal{K}_u$. Then $BF \in \mathcal{K}_u$ if and only if B divides i_F.*

Proof Suppose that $f = BF \in \mathcal{K}_u$. Since $Cf = \overline{fz}u = \overline{BFz}u$ on $\mathbb{T}$, (13.2.3) implies that

$$BCf = \overline{Fz}u = \mathrm{i}_F F$$

on $\mathbb{T}$. The identity principle guarantees that $BCf = \mathrm{i}_F F$ on $\mathbb{D}$ and hence the uniqueness of the inner–outer factorization confirms that B divides i_F.

Conversely, suppose that B divides i_F. Then $\mathrm{i}_F = BB_1$, in which B_1 is a finite Blaschke product. From (13.2.3) we see that $BB_1F = \overline{F}u$ on $\mathbb{T}$ and hence

$$BF = \overline{B_1Fz}u.$$

Proposition 12.5.5 shows that $BF \in \mathcal{K}_u$ and that $C(BF) = B_1F$. □

Example 13.2.5 The constant function 1 belongs to $\mathcal{K}_{zu}$ (Proposition 12.1.9) and $C1 = \overline{1z}zu = u$. In this case, $F \equiv 1$ and $\mathrm{i}_F = u$. Thus, the only finite Blaschke products that belong to $\mathcal{K}_{zu}$ are the divisors of u.

Example 13.2.6 Building upon Example 13.1.5, Proposition 13.2.4 implies that the only finite Blaschke products B for which $Bk_\lambda \in \mathcal{K}_u$ are the divisors of the finite Blaschke product $(b_{u(\lambda)} \circ u)/b_\lambda$.

13.3 Quaternionic Structure of Solutions

Let $a \in H^\infty$ be a rational function and suppose that the corresponding Darlington synthesis problem has a solution U, in which b, c, and d are rational functions in H^∞. Then the matrix-valued function U from (13.0.1) is analytic on $\mathbb{D}$ and unitary on $\mathbb{T}$ and hence $u = \det U$ is a finite Blaschke product. Indeed,

$$\begin{aligned}
|u(\zeta)|^2 &= |\det U(\zeta)|^2 \\
&= \overline{\det U(\zeta)}\det U(\zeta) \\
&= \det U^*(\zeta)\det U(\zeta) \\
&= \det U^*(\zeta)U(\zeta) \\
&= \det I \\
&= 1
\end{aligned}$$

for $\zeta \in \mathbb{T}$, so Theorem 3.5.2 implies that u is a finite Blaschke product.

The precise relationship between the finite Blaschke product u and the entries of U is given in the following theorem from [51], where it is proved without the assumption that a is rational (the proof is largely similar).

Theorem 13.3.1 *Let $a, b, c, d \in H^\infty$ be rational functions and let u be a finite Blaschke product. Then the matrix-valued analytic function U given by (13.0.1) is unitary on $\mathbb{T}$ and satisfies* $\det U = u$ *if and only if the following hold.*

(a) $a, b, c, d \in \mathscr{K}_{zu}$.
(b) $Ca = d$ *and* $Cb = c$*, where C is the conjugation on $\mathscr{K}_{zu}$; that is,* $Cf = \overline{f}u$ *on* $\mathbb{T}$.
(c) $|a|^2 + |b|^2 = 1$ *on* $\mathbb{T}$.

Proof ($\Rightarrow$) If U is unitary on $\mathbb{T}$, then $u = \det U$ is a finite Blaschke product by the argument above. Compare entries in $U = (U^*)^{-1}$ and obtain

$$a = \overline{d}u \qquad \text{and} \qquad b = \overline{c}u$$

on $\mathbb{T}$. Proposition 12.5.5 implies that $a, b, c, d \in \mathscr{K}_{zu}$, $Ca = d$, and $Cb = c$. Examine the diagonal entries of the identity $UU^* = I$ to obtain $|a|^2 + |b|^2 = 1$ on $\mathbb{T}$.

($\Leftarrow$) Suppose that (a), (b), and (c) hold. Write $a = B_a F$, $b = B_b G$, $c = B_c G$, and $d = B_d F$, in which B_a, B_b, B_c, B_d are finite Blaschke products and F, G are rational outer functions in $\mathscr{K}_{zu}$. Observe that Proposition 12.5.4 and Lemma 13.1.4 say that the outer factors of a and $d = Ca$ are the same (as are the outer factors of b and $c = Cb$). Condition (c) implies that the entries on the main diagonal of the matrix product

$$UU^* = \begin{bmatrix} B_a F & -B_b G \\ B_c G & B_d F \end{bmatrix} \begin{bmatrix} \overline{B_a F} & \overline{B_c G} \\ \overline{-B_b G} & \overline{B_d F} \end{bmatrix}. \tag{13.3.2}$$

are both identically 1 on $\mathbb{T}$. The upper right-hand corner of the product (13.3.2) is

$$X = B_a F \overline{B_c G} - B_b G \overline{B_d F},$$

and a few more manipulations yield

$$X \frac{B_c B_d}{\overline{FG}} = B_a B_d \frac{F}{\overline{F}} - B_b B_c \frac{G}{\overline{G}}$$

on $\mathbb{T}$ (except for at most finitely many poles). Since a, d and b, c are pairs of conjugates in $\mathscr{K}_{zu}$, it follows from (13.2.2) that $B_a B_d = \mathrm{i}_F$ and $B_b B_c = \mathrm{i}_G$. Then

$$X \frac{B_c B_d}{\overline{FG}} = \mathrm{i}_F \frac{F}{\overline{F}} - \mathrm{i}_G \frac{G}{\overline{G}} = u - u = 0$$

on $\mathbb{T}$ by (13.2.3). The identity principle implies that X vanishes identically. A similar argument shows that the bottom left-hand corner of the matrix product (13.3.2) vanishes and hence U is unitary on $\mathbb{T}$. To complete the proof we use (13.2.3) to compute:

$$\begin{aligned}\det U &= ad + bc \\ &= B_a B_d F^2 + B_b B_c G^2 \\ &= \mathrm{i}_F F^2 + \mathrm{i}_G G^2 \\ &= |F|^2 u + |G|^2 u \\ &= u.\end{aligned}$$

This completes the proof. □

As a byproduct of Theorem 13.3.1 we obtain two convenient representations for $u = \det U$:

$$u = a\,Ca + b\,Cb \tag{13.3.3}$$

$$= \mathrm{i}_F F^2 + \mathrm{i}_G G^2. \tag{13.3.4}$$

These formulas will be useful in what follows.

Theorem 13.3.1 shows that any solution to the Darlington synthesis problem is (almost) a quaternion-valued function on $\mathbb{T}$ with values of unit modulus on $\mathbb{T}$. Recall that the *quaternions* are a division algebra, denoted by $\mathbb{H}$ in honor of their discoverer William Rowan Hamilton, that consists of all expressions of the form

$$\alpha + \beta i + \delta j + \gamma k,$$

in which $\alpha, \beta, \delta, \gamma \in \mathbb{R}$ and i, j, and k are symbols that satisfy

$$i^2 = j^2 = k^2 = -1, \qquad ij = k, \quad \text{and} \quad ji = -k.$$

One can show that $\mathbb{H}$ is isomorphic to the division algebra formed by the complex matrices

$$\begin{bmatrix} z & -w \\ \overline{w} & \overline{z} \end{bmatrix}, \tag{13.3.5}$$

in which

$$z = \alpha + \beta i \quad \text{and} \quad w = -(\delta + i\gamma).$$

The square of the absolute value of the quaternion $\alpha + \beta i + \delta j + \gamma k$ is

$$|z|^2 + |w|^2 = \alpha^2 + \beta^2 + \delta^2 + \gamma^2.$$

The matrix representation (13.3.5) is reminiscent of that obtained in Theorem 13.3.1, which asserts that any solution to the Darlington synthesis problem with data a is of the form

$$U = \begin{bmatrix} a & -b \\ Cb & Ca \end{bmatrix} = \begin{bmatrix} a & -b \\ \overline{b}u & \overline{a}u \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & u \end{bmatrix} \begin{bmatrix} a & -b \\ \overline{b} & \overline{a} \end{bmatrix}.$$

13.4 Primitive Solution Sets

Suppose that the scalar-valued Darlington synthesis problem with rational data $a \in H^\infty$ is solvable and that U is one particular solution. That is, U is a 2×2 rational, matrix-valued analytic function on $\mathbb{D}$ and its $(1, 1)$ entry is the function a. Theorem 13.3.1 tells us that $u = \det U$ is a finite Blaschke product and

$$U = \begin{bmatrix} a & -b \\ Cb & Ca \end{bmatrix},$$

in which Ca and Cb are the respective conjugates of a and b in $\mathscr{K}_{zu}$. Each such U provides us with infinitely many other rational solutions via the following method.

If B_1 and B_2 are finite Blaschke products (possibly constant), then

$$V = \begin{bmatrix} a & -B_1 b \\ B_2 Cb & B_1 B_2 Ca \end{bmatrix} \tag{13.4.1}$$

is another solution to the Darlington synthesis problem with data a. Indeed, V is analytic on $\mathbb{D}$ and satisfies

$$\begin{aligned} V^* V &= \begin{bmatrix} \overline{a} & \overline{B_2 Cb} \\ \overline{-B_1 b} & \overline{B_1 B_2 Ca} \end{bmatrix} \begin{bmatrix} a & -B_1 b \\ B_2 Cb & B_1 B_2 Ca \end{bmatrix} \\ &= \begin{bmatrix} |a|^2 + |Cb|^2 & -\overline{a} B_1 b + B_1 Ca \overline{Cb} \\ -a \overline{B_1 b} + \overline{B_1 Ca} Cb & -|b|^2 + |Ca|^2 \end{bmatrix} \\ &= \begin{bmatrix} |a|^2 + |b|^2 & B_1(-\overline{a} b + \overline{a} u b \overline{u}) \\ \overline{B_1}(-a \overline{b} + a \overline{u} \overline{b} u) & |a|^2 + |b|^2 \end{bmatrix} \\ &= I. \end{aligned}$$

As predicted by Theorem 13.3.1, we may write

$$V = \begin{bmatrix} a & -(B_1 b) \\ C_v(B_1 b) & C_v a \end{bmatrix},$$

in which

$$v = \det V = B_1 B_2 u$$

is a finite Blaschke product and

$$C_v f = \overline{f}(B_1 B_2 u)$$

is the conjugation on $\mathcal{K}_{zB_1B_2u}$. Since u divides v, we are prompted to consider solutions to the Darlington synthesis problem that have minimal determinant.

We say that a solution U is *primitive* if the finite Blaschke product $u = \det U$ is the minimal finite Blaschke such that $\det U$ divides $\det V$ for any other solution V. This is equivalent to requiring that u is the minimal finite Blaschke product such that a belongs to $\mathcal{K}_{zu}$. Note also that every primitive solution shares the same determinant, up to a unimodular constant factor. We call u the *minimal determinant* for the problem.

Any solution V to the Darlington synthesis problem with data a can be written in terms of a primitive solution via (13.4.1). Indeed, suppose that V is a solution with determinant $V = uv$, in which v is a finite Blaschke product and u is the minimal determinant for the problem. Let the outer functions F and G be defined as in the proof of Theorem 13.3.1 and let i_F and i_G denote the finite Blaschke products associated with F and G with respect to u. That is,

$$\mathrm{i}_F F = \overline{F} u \qquad \text{and} \qquad \mathrm{i}_G G = \overline{G} u$$

on $\mathbb{T}$. Theorem 13.3.1 permits us to write

$$V = \begin{bmatrix} a & -c \\ \overline{c}uv & \overline{a}uv \end{bmatrix},$$

in which $c \in \mathcal{K}_{zuv}$ and has outer factor G. Since the conjugate Ca of a in $\mathcal{K}_{zu}$ equals $\overline{a}u$ on $\mathbb{T}$, we conclude that

$$\det V = uv = \mathrm{i}_F F^2 v + BG^2 v,$$

in which B is a finite Blaschke product. Compare this with (13.3.4) to conclude that $\mathrm{i}_G = B$. In particular,

$$c(\overline{c}uv) = \mathrm{i}_G v G^2$$

since on $\mathbb{T}$, $\overline{c}uv$ is the boundary function for the conjugate of c in $\mathcal{K}_{zuv}$. Thus, V can be written in the form (13.4.1) for some finite Blaschke products B_1 and B_2 such that $B_1B_2 = v$.

It therefore suffices to describe all primitive solutions. We call a complete collection of primitive solutions sharing the same minimal determinant a *primitive solution set*. Since the minimal determinant is determined only up to a unimodular constant factor, there will be infinitely many primitive solution sets. These are related to each other by (13.4.1), in which the finite Blaschke products B_1 and B_2 are just unimodular constants.

Fix a minimal determinant u to the Darlington synthesis problem with data a; that is, u is minimal with the property that $a \in \mathcal{K}_{zu}$. We want to describe all solutions U with $\det U = u$. Theorem 13.3.1 permits us to identify each such solution with its upper right-hand entry, b. Indeed, a, b, and u uniquely determine the remaining entries $Cb = \overline{b}u$ and $Ca = \overline{a}u$. Since the outer factor G of b is determined by condition (c) of Theorem 13.3.1, we can identify each solution U with the maximal finite Blaschke product that divides b.

Since b divides i_G, which is determined by (13.3.4), there is a bijective correspondence between elements of our primitive solution set and the finite Blaschke products that divide i_G. In particular, a primitive solution set has a natural partial order that derives from this correspondence.

Example 13.4.2 If i_G is a unimodular constant, then each primitive solution set consists of precisely one solution. Since i_G is the product of the finite Blaschke product factors of b and Cb, this occurs precisely when b is a self-conjugate, rational outer function. For example, suppose that u is a finite Blaschke product and

$$a = \tfrac{1}{2}(1 + u). \tag{13.4.3}$$

Since a generates $\mathcal{K}_{zu}$ (Exercise 13.1), any solution to the Darlington synthesis problem with data a is primitive. Since

$$Ca = \tfrac{1}{2}(C1 + Cu) = \tfrac{1}{2}(u + 1) = a,$$

we see that a is self-conjugate. Moreover, the rational outer function

$$b = \tfrac{1}{2i}(1 - u)$$

belongs to $\mathcal{K}_{zu}$ and is also self-conjugate. In particular, $\mathrm{i}_G = 1$. Thus,

$$U = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

is the unique solution to the Darlington synthesis problem with data a that has minimal determinant u.

Example 13.4.4 Suppose that i_G is a finite Blaschke product of order n. Then a primitive solution set contains at most 2^n solutions. The exact number depends upon the multiplicities of the zeros of i_G. Moreover, a primitive solution set is linearly ordered (via divisibility of the corresponding inner functions) if and only if i_G is a unimodular multiple of a power of a single Blaschke factor.

Example 13.4.5 If i_G is the square of a finite Blaschke product, then a symmetric ($U = U^{\mathsf{T}}$) primitive solution exists. If $\mathrm{i}_G = B^2$, in which B is a finite Blaschke product, then $b = BG$ belongs to $\mathscr{K}_{zu}$ and is self-conjugate. This yields the primitive solution

$$\begin{bmatrix} a & -b \\ b & Ca \end{bmatrix}.$$

Using (13.4.1) with $B_1 = -i$ and $B_2 = i$, we obtain the symmetric solution

$$\begin{bmatrix} a & ib \\ ib & Ca \end{bmatrix}.$$

13.5 Construction of the Solutions

Suppose that $a \in H^\infty$ is a rational function with $\|a\|_\infty \leqslant 1$. Write $a = P/R$, in which P is a polynomial that is relatively prime to

$$R(z) = (1 - \overline{\lambda_1} z)(1 - \overline{\lambda_2} z) \cdots (1 - \overline{\lambda_n} z).$$

Since a is bounded on $\mathbb{D}^-$, it follows that $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{D}$. There are two cases to consider, depending on $m = \deg P$.

Case I If $m \leqslant n$, then $a \in \mathscr{K}_{zu}$, where

$$u(z) = \prod_{k=1}^{n} \frac{z - \lambda_k}{1 - \overline{\lambda_k} z}.$$

By (12.1.2) every function in $\mathscr{K}_{zu}$ is of the form Q/R, in which Q is a polynomial of degree at most n. Since u is the minimal (up to a unimodular constant factor) finite Blaschke product so that $a \in \mathscr{K}_{zu}$, to find all primitive solutions to the Darlington synthesis problem with data a, it suffices to describe all solutions U with $u = \det U$.

By the definition of the conjugation C from (12.5.3),

$$C(Q/R) = \frac{Q^\#}{R}, \tag{13.5.1}$$

in which

$$Q^{\#}(z) = z^n \overline{Q(1/\overline{z})}.$$

We require two special cases of (13.5.1):

$$Ca = P^{\#}/R \qquad \text{and} \qquad u = R^{\#}/R. \tag{13.5.2}$$

The second formula follows from the fact that $C1 = C(R/R) = R^{\#}/R$; that is, 1 and u are conjugates in $\mathcal{K}_{zu}$.

By Theorem 13.3.1 and (13.3.3), the desired U are of the form

$$U = \begin{bmatrix} a & -b \\ Cb & Ca \end{bmatrix},$$

in which

$$u = a\,Ca + b\,Cb. \tag{13.5.3}$$

Write $b = Q/R$, in which Q is a polynomial of degree at most n. In light of (13.5.1) and (13.5.2), to solve (13.5.3) we must solve

$$\frac{R^{\#}}{R} = \frac{P^{\#}P}{R^2} + \frac{Q^{\#}Q}{R^2} \tag{13.5.4}$$

for Q. This reduces to

$$Q^{\#}Q = R^{\#}R - P^{\#}P. \tag{13.5.5}$$

Note that $P^{\#}$ and $R^{\#}$ can be obtained from a without factoring R. Consequently, one can proceed directly to (13.5.5).

Write $b = B_b G$ and $Cb = B_{Cb}G$, in which B_b and B_{Cb} are finite Blaschke products and G is the common outer factor of b and Cb. Since

$$b\,Cb = \frac{R^{\#}R - P^{\#}P}{R^2},$$

it follows that

$$\mathrm{i}_G G^2 = B_b B_{Cb} G^2 = \frac{R^{\#}R - P^{\#}P}{R^2}. \tag{13.5.6}$$

As discussed in Sect. 13.4, we need only find G and i_G in order to parameterize all solutions U with $\det U = u$. To find these functions, we need only produce the inner–outer factorization of

$$\frac{R^{\#}R - P^{\#}P}{R^2},$$

a rational function that is directly obtained from the datum a.

Since the outer factor of any function in $\mathscr{K}_{zu}$ also lies in $\mathscr{K}_{zu}$, it follows that $G \in \mathscr{K}_{zu}$. Therefore, $G = S/R$, in which S is a polynomial of degree at most n. Since G and R are outer, it follows that S is outer. Thus, (13.5.6) reduces to

$$\mathrm{i}_G S^2 = R^{\#}R - P^{\#}P, \tag{13.5.7}$$

in which i_G is a finite Blaschke product, possibly constant, whose zeros are the zeros of $R^{\#}R - P^{\#}P$ that lie in $\mathbb{D}$, repeated according to multiplicity. Although the degree of $R^{\#}R - P^{\#}P$ is $2n$, at most n of its zeros (up to multiplicity) belong to $\mathbb{D}$. This is because $R^{\#}R - P^{\#}P$ is invariant under the transformation $z \mapsto 1/\overline{z}$. This yields the (possibly identical) solutions

$$\begin{bmatrix} P/R & -S/R \\ S^{\#}/R & P^{\#}/R \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} P/R & -S^{\#}/R \\ S/R & P^{\#}/R \end{bmatrix}.$$

We now identify the remaining (if any) primitive solutions with determinant u. Since $G = S/R$ is an outer function in $\mathscr{K}_{zu}$, we have

$$\widehat{G} = \mathrm{i}_G G = \frac{S^{\#}}{R}.$$

Therefore the desired finite Blaschke product i_G is given by

$$\mathrm{i}_G = \frac{S^{\#}}{S}.$$

Since S is an outer function, the zeros of i_G must be precisely the zeros of $S^{\#}$ that lie in $\mathbb{D}$. However, S and $S^{\#}$ may have common zeros on $\mathbb{T}$. We can discard these without actually finding them by simply calculating the greatest common divisor of the polynomials S and $S^{\#}$ (this can be accomplished using the Euclidean algorithm and hence it does not require factoring S or $S^{\#}$). Without loss of generality, we assume that the zeros of $S^{\#}$ all lie in $\mathbb{D}$.

Once the zeros of $S^{\#}$ have been found, the primitive solutions with determinant u can be identified with functions

$$b = B_b G = B_b \frac{S}{R},$$

in which B_b is a finite Blaschke product that divides i_G. The polynomials Q from (13.5.4) are the functions $B_b S$.

Case II If $m > n$, then use $z^{m-n}u$ in place of u and define

$$Q^{\#}(z) = z^m \overline{Q(1/\overline{z})}$$

for polynomials Q of degree at most m. We may proceed as before, the only difference being that $R^{\#}R - P^{\#}P$ is now of degree at most $2m$. The details are left to the reader; see Exercise 13.2.

What is the significance of the polynomial $R^{\#}R - P^{\#}P$? For the sake of simplicity, we suppose that $m \leqslant n$. Since

$$\frac{R^{\#}R - P^{\#}P}{R^2} = u - a\,Ca,$$

the zeros of $R^{\#}R - P^{\#}P$ correspond are the zeros of $u - a\,Ca$. On $\mathbb{T}$, we have

$$\frac{R^{\#}R - P^{\#}P}{R^2} = u(1 - |a|^2).$$

Consequently, the zeros of $R^{\#}R - P^{\#}P$ that lie on $\mathbb{T}$ are exactly the points at which $|a| = 1$. Since the zeros of $R^{\#}R - P^{\#}P$ occur in pairs symmetric with respect to $\mathbb{T}$, the number of zeros inside the unit disk, counted according to multiplicity, depends on the degree of $R^{\#}R - P^{\#}P$ and the number of times, according to multiplicity, that a assumes its maximum modulus on $\mathbb{T}$. Thus, the number of solutions in a primitive solution set depends on how many times the datum a assumes values on $\mathbb{T}$.

Example 13.5.8 If $\deg(R^{\#}R - P^{\#}P) = 2n$ and a assumes values with modulus one n times on $\mathbb{T}$, then $R^{\#}R - P^{\#}P$, and hence i_G, has no zeros in $\mathbb{D}$. In this case, the solution Darlington synthesis problem with data a is essentially unique because each primitive solution set contains only one solution; see Example 13.4.2.

We conclude with an algorithm to produce a complete primitive solution set to the scalar valued Darlington synthesis problem.

Algorithm

Suppose that we are given a rational $a \in H^\infty$ with $\|a\|_\infty \leqslant 1$.

(a) Write $a = P/R$, in which R is a polynomial with constant term 1 and P is relatively prime to R. Let $m = \deg P$ and $n = \deg R$.
(b) If $m \leqslant n$, then form the polynomial $R^{\#}R - P^{\#}P$, in which $Q^{\#}(z) = z^n\overline{Q(1/\overline{z})}$ for polynomials $Q(z)$ of degree at most n.

   (i) The outer factor of $R^{\#}R - P^{\#}P$ is a polynomial S^2 of degree at most $2n$ (see (13.5.7)). Then

$$\begin{bmatrix} P/R & -S/R \\ S^{\#}/R & P^{\#}/R \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} P/R & -S^{\#}/R \\ S/R & P^{\#}/R \end{bmatrix}$$

   are primitive solutions with determinant $u = R^{\#}/R$.

(ii) The roots of the polynomial

$$S_1 = \frac{S^\#}{\gcd(S, S^\#)},$$

which is of degree $N \leqslant n$, all lie in $\mathbb{D}$.

(iii) For each subset $\{z_1, z_2, \ldots, z_k\}$ of the roots of S_1 such that $k \leqslant \lfloor \frac{N}{2} \rfloor$,

$$T(z) = S(z) \prod_{j=1}^{k} \frac{z - z_j}{1 - \overline{z_j} z}$$

is a polynomial of degree $N - k$ that yields the primitive solutions

$$\begin{bmatrix} P/R & -T/R \\ T^\#/R & P^\#/R \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} P/R & -T^\#/R \\ T/R & P^\#/R \end{bmatrix}.$$

This yields a complete set of primitive solutions with determinant u.

(c) If $m > n$, then form the polynomial $R^\# R - P^\# P$ (of degree at most $2m$) using the definition $Q^\#(z) = z^m \overline{Q(1/\overline{z})}$ for polynomials Q of degree at most m. Proceed as in the previous case.

13.6 Notes

The Schur–Cohn algorithm [91] can detect the number of zeros of a polynomial inside the disk, on its boundary, and outside. Therefore in many situations, we can produce information on the number of solutions in a primitive solution set without explicitly finding the roots of polynomials.

13.7 Exercises

13.1 Show that the function $a = \frac{1}{2}(1 + u)$ generates $\mathscr{K}_{zu}$; that is,

$$\operatorname{span}\{S^{*n} a : n = 0, 1, 2, \ldots\} = \mathscr{K}_{zu},$$

in which S^* is the backward shift operator.

13.2 Fill in the details of the construction from Sect. 13.5.

13.3 Example 13.1.3 suggests a method for factoring

$$f(z) = \frac{p(z)}{\prod_{j=1}^{n}(1 - \overline{\lambda_j} z)} \in \mathscr{K}_u$$

into its inner and outer factors. Assume that f is not identically zero and write

$$p(z) = c\prod_{r=1}^{R}(z-\lambda_r)\cdot\prod_{s=1}^{S}(z-\zeta_s)\cdot\prod_{t=1}^{T}(z-w_t), \tag{13.7.1}$$

where

$$c \neq 0, \quad \lambda_r \in \mathbb{D}, \quad \zeta_s \in \mathbb{T}, \quad w_t \in \mathbb{C}\backslash\mathbb{D}^-, \quad \text{and} \quad R+S+T = \deg p \leqslant n-1.$$

(a) Divide f by the finite Blaschke product

$$B_f(z) = \prod_{r=1}^{R}\frac{z-\lambda_r}{1-\overline{\lambda_r}z}$$

and verify that

$$F(z) = c\prod_{s=1}^{S}(z-\zeta_s)\cdot\prod_{t=1}^{T}(z-w_t)$$

is the outer factor of f.

(b) Verify that F belongs to $\mathscr{K}_u$.

13.4 If

$$f(z) = \frac{p(z)}{\prod_{j=1}^{n}(1-\overline{\lambda_j}z)} \in \mathscr{K}_u,$$

with

$$p(z) = c\prod_{r=1}^{R}(z-\zeta_r)\cdot\prod_{s=1}^{S}(z-\lambda_s)\cdot\prod_{t=1}^{T}(z-w_t)$$

as in (13.7.1), prove that

$$B_f(z) = \prod_{s=1}^{S}\frac{z-\lambda_s}{1-\overline{\lambda_s}z} \qquad \text{and} \qquad B_{Cf}(z) = \prod_{t=1}^{T}\frac{z-1/\overline{w_t}}{1-z/w_t},$$

from which it follows that

$$\mathrm{i}_F = B_f B_{Cf} = \prod_{s=1}^{S}\frac{z-\lambda_s}{1-\overline{\lambda_s}z}\cdot\prod_{t=1}^{T}\frac{z-1/\overline{w_t}}{1-z/w_t}.$$

Appendix A
Some Reminders

This appendix briefly covers several peripheral topics that have arisen occasionally in the preceding text. We do not aim to give a complete account of any of the following subjects and we provide relatively few proofs. The reader is invited to consult the references discussed below for more details.

A.1 Fourier Analysis

Let $L^2 := L^2(\mathbb{T}, d\theta/2\pi)$ denote the space of complex-valued Lebesgue measurable functions on $\mathbb{T}$ such that

$$\|f\|_{L^2} := \sqrt{\int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi}}$$

is finite. The Lebesgue theory of integration [122, 123] can be used to show that L^2 is a Hilbert space when endowed with inner product

$$\langle f, g\rangle_{L^2} = \int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})}\,\frac{d\theta}{2\pi} \tag{A.1.1}$$

arising from the norm $\|\cdot\|_{L^2}$. For $f \in L^2$ and $n \in \mathbb{Z}$, the *nth Fourier coefficient* of f is

$$\widehat{f}(n) := \int_0^{2\pi} f(e^{i\theta})e^{-in\theta}\,\frac{d\theta}{2\pi} \tag{A.1.2}$$

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8

and the *Fourier series* of f is

$$\sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{in\theta}.$$

Convergence of a Fourier series is with respect to the L^2-norm.

Theorem A.1.3 (Parseval's Theorem) *For* $f \in L^2$, $\|f\|_{L^2}^2 = \sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2$.

A.2 The Cauchy Integral Formula

The Cauchy integral formula [123] says that if f is analytic on a neighborhood of $\mathbb{D}^-$ and $\lambda \in \mathbb{D}$, then

$$f(\lambda) = \frac{1}{2\pi i} \oint_{\mathbb{T}} \frac{f(\xi)}{\xi - \lambda} d\xi. \tag{A.2.1}$$

We often use the following rephrasing in terms of L^2-inner products.

Lemma A.2.2 (Cauchy Integral Formula) *Suppose that* f *is analytic on a neighborhood of* $\mathbb{D}^-$ *and* $\lambda \in \mathbb{D}$*. Then* $f|_{\mathbb{T}} \in L^2$ *and*

$$\left\langle f, \frac{1}{1-\overline{\lambda} z} \right\rangle = f(\lambda).$$

Proof By the definition of the L^2-inner product from (A.1.1) we have

$$\begin{aligned}
\left\langle f, \frac{1}{1-\overline{\lambda} z} \right\rangle &= \int_0^{2\pi} f(e^{i\theta}) \overline{\frac{1}{1-\overline{\lambda} e^{i\theta}}} \frac{d\theta}{2\pi} \\
&= \int_0^{2\pi} f(e^{i\theta}) \frac{1}{1-\lambda e^{-i\theta}} \frac{d\theta}{2\pi} \\
&= \int_0^{2\pi} f(e^{i\theta}) \frac{e^{i\theta}}{e^{i\theta}-\lambda} \frac{d\theta}{2\pi} \\
&= \frac{1}{2\pi i} \oint_{\mathbb{T}} \frac{f(\xi)}{\xi - \lambda} d\xi \qquad (\xi = e^{i\theta}) \\
&= f(\lambda).
\end{aligned}$$

The final equality is due to the Cauchy integral formula from (A.2.1). ◻

A.3 Fatou's Theorem

Analytic functions on $\mathbb{D}$ need not have limiting values anywhere on $\mathbb{T}$. However, bounded analytic functions are nicer. This is made more precise with the following theorem of Fatou [38].

Theorem A.3.1 (Fatou) *If f is a bounded analytic function on $\mathbb{D}$, then the radial limit*

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

exists for almost every $\theta \in [0, 2\pi]$.

A.4 Hardy Space Theory

Here are some standard facts about the Hardy space H^2. Several good sources are [38, 61, 101]. The *Hardy space H^2* is the set of analytic functions f on $\mathbb{D}$ for which

$$\sup_{0<r<1} \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi} < \infty.$$

A well-known result from H^2 theory says that for almost every $\theta \in [0, 2\pi]$, the radial limit

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

exists (and is finite). Furthermore, the boundary function $e^{i\theta} \mapsto f(e^{i\theta})$ belongs to L^2 and satisfies

$$\int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi} = \sup_{0<r<1} \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi}. \tag{A.4.1}$$

Thus, via radial limits and boundary functions, one can view H^2 as a closed subspace of L^2. If $f \in H^2$ has the Taylor series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

about $z = 0$, then

$$a_n = \int_0^{2\pi} f(e^{i\theta}) e^{-in\theta} \frac{d\theta}{2\pi}.$$

In other words, the Fourier coefficients of the (almost everywhere defined) boundary function $e^{i\theta} \mapsto f(e^{i\theta})$ are equal to the corresponding Taylor coefficients. From here one can view H^2 as

$$H^2 = \{f \in L^2 : \widehat{f}(n) = 0 \text{ for } n \leqslant -1\}.$$

The *Riesz projection* is the operator $P : L^2 \to L^2$ with range equal to H^2 defined by

$$P\Big(\sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{in\theta} \Big) := \sum_{n=0}^{\infty} \widehat{f}(n) e^{in\theta}. \tag{A.4.2}$$

We also have the following *Parseval's formula* for H^2 functions:

$$\sum_{n=0}^{\infty} |a_n|^2 = \int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi}. \tag{A.4.3}$$

The integral computation

$$\int_0^{2\pi} e^{in\theta} \frac{d\theta}{2\pi} = \begin{cases} 0 & \text{if } n \neq 0, \\ 1 & \text{if } n = 0, \end{cases} \tag{A.4.4}$$

shows that $\{1, z, z^2, \ldots\}$ is an orthonormal set in H^2. In fact, it is an orthonormal basis for H^2.

From Parseval's formula, we obtain the identities

$$\sum_{n=0}^{\infty} a_n \overline{b_n} = \int_0^{2\pi} f(e^{i\theta}) \overline{g(e^{i\theta})} \frac{d\theta}{2\pi} \tag{A.4.5}$$

and

$$\sum_{n=0}^{\infty} r^{2n} |a_n|^2 = \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi}. \tag{A.4.6}$$

A.5 Jensen's Formula and Jensen's Inequality

Suppose that f is an analytic function on $|z| \leqslant r$ with no zeros. The mean value property for harmonic functions, applied to the harmonic function $\log|f|$, says that

$$\log|f(0)| = \int_0^{2\pi} \log|f(re^{i\theta})| \frac{d\theta}{2\pi}.$$

For a function with finitely many zeros $z_1, z_2, \ldots, z_n$, there is the following generalization [123].

Theorem A.5.1 (Jensen's Formula) *Let f be an analytic function on $|z| \leqslant r$ with $f(0) \neq 0$ and with zeros $z_1, z_2, \ldots, z_n$ in $|z| < r$. Then*

$$\log|f(0)| = \sum_{k=1}^{n} \log\left(\frac{|z_k|}{r}\right) + \int_0^{2\pi} \log|f(re^{i\theta})| \frac{d\theta}{2\pi}.$$

There is a probabilistic inequality, also due to Jensen, which says the following.

Theorem A.5.2 (Jensen's Inequality) *Suppose that (Ω, A, μ) is a probability space and g is a real-valued μ-integrable function on Ω. If ϕ is a real-valued convex function on $\mathbb{R}$, then*

$$\phi\left(\int_\Omega g \, d\mu\right) \leqslant \int_\Omega \phi \circ g \, d\mu.$$

A.6 Hilbert Spaces and Their Operators

An excellent source for operators on Hilbert spaces is [27]. The proofs of the material presented below can be found there.

Inner Product

Let $\mathscr{H}$ be a separable complex Hilbert space with inner product $\langle \mathbf{x}, \mathbf{y} \rangle$ and corresponding norm $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. Three important facts that will be used in this book are the *polarization identity*

$$4\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 + i\|\mathbf{x} + i\mathbf{y}\|^2 - i\|\mathbf{x} - i\mathbf{y}\|^2,$$

the *parallelogram identity*

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2),$$

and the *Cauchy–Schwarz inequality*

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\|.$$

The Vector Space $\mathbb{C}^n$

Finite-dimensional Hilbert spaces play a prominent role in this book. One of the most important finite-dimensional Hilbert spaces is the complex vector space

$$\mathbb{C}^n := \{(x_1, x_2, \ldots, x_n) : x_j \in \mathbb{C}\}. \tag{A.6.1}$$

We will use the notation

$$\mathbf{e}_j = (0, 0, \ldots, 0, 1, 0, 0, \ldots, 0),$$

where the 1 appears in the jth slot. Observe that $\{\mathbf{e}_j : j = 1, 2, \ldots, n\}$ is a basis for $\mathbb{C}^n$, called the *standard basis*. We can make $\mathbb{C}^n$ into a Hilbert space if we endow it with the inner product

$$\langle \mathbf{x}, \mathbf{y}\rangle_{\mathbb{C}^n} := \sum_{j=1}^{n} x_j \overline{y_j}.$$

This yields the norm

$$\|\mathbf{x}\|_{\mathbb{C}^n} = \sqrt{\sum_{j=1}^{n} |x_j|^2}.$$

Notice that $\{\mathbf{e}_j : j = 1, 2, \ldots, n\}$ is an orthonormal basis for $\mathbb{C}^n$.

Operators

Let $\mathscr{H}$ be a separable Hilbert space. A linear transformation (operator) $T : \mathscr{H} \to \mathscr{H}$ is *bounded* if

$$\sup\{\|T\mathbf{x}\| : \mathbf{x} \in \mathscr{H}, \|\mathbf{x}\| \leqslant 1\} < \infty.$$

The *operator norm* $\|T\|$ of T is

$$\|T\| := \sup\{\|T\mathbf{x}\| : \mathbf{x} \in \mathscr{H}, \|\mathbf{x}\| \leqslant 1\}. \tag{A.6.2}$$

The set of all bounded linear operators on $\mathscr{H}$ is denoted $\mathscr{L}(\mathscr{H})$. Observe that $\mathscr{L}(\mathscr{H})$ is a complex linear space as well as a normed algebra whose norm is submultiplicative:

$$\|T_1T_2\| \leqslant \|T_1\|\|T_2\|, \quad T_1, T_2 \in \mathscr{L}(\mathscr{H}).$$

For $T \in \mathscr{L}(\mathscr{H})$, the *adjoint* of T, denoted by T^*, is the unique $T^* \in \mathscr{L}(\mathscr{H})$ satisfying

$$\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, T^*\mathbf{y}\rangle, \quad \mathbf{x}, \mathbf{y} \in \mathscr{H}.$$

One can show that $(T^*)^* = T$,

$$\|T\| = \|T^*\|,$$

and

$$\|T^*T\| = \|T\|^2. \tag{A.6.3}$$

With respect to the operator norm, $\mathscr{L}(\mathscr{H})$ is a C^**-algebra*; that is, a complete normed algebra with an involution $*$ that satisfies (A.6.3).

Definition A.6.4 $T \in \mathscr{L}(\mathscr{H})$ is

(a) *self-adjoint* if $T = T^*$;
(b) *unitary* if $TT^* = T^*T = I$;
(c) *positive semidefinite* if $\langle T\mathbf{x}, \mathbf{x}\rangle \geqslant 0$ for all $\mathbf{x} \in \mathscr{H}$;
(d) a *contraction* if $\|T\| \leqslant 1$.

If T is positive semidefinite, then we write $T \geqslant 0$. The notation $T \leqslant S$ indicates that $S - T$ is positive semidefinite. A short exercise shows that

$$I - T^*T \geqslant 0 \qquad \Longleftrightarrow \qquad \|T\| \leqslant 1. \tag{A.6.5}$$

If T is a strict contraction, that is $\|T\| < 1$, then

$$\sum_{n=0}^{\infty} T^n = (I - T)^{-1}. \tag{A.6.6}$$

Convergence of the series above is with respect to the operator norm.

Orthogonal Projections

If $\mathscr{M}$ is a closed subspace of $\mathscr{H}$, the *orthogonal complement* of $\mathscr{M}$ is

$$\mathscr{M}^{\perp} := \{\mathbf{x} \in \mathscr{H} : \langle \mathbf{x}, \mathbf{y} \rangle = 0 \ \ \forall \mathbf{y} \in \mathscr{M}\}.$$

Note that $\mathscr{M}^{\perp}$ is a closed subspace of $\mathscr{H}$ and that every $\mathbf{x} \in \mathscr{H}$ can be written uniquely as

$$\mathbf{x} = \mathbf{x}_{\mathscr{M}} + \mathbf{x}_{\mathscr{M}^{\perp}}, \quad \mathbf{x}_{\mathscr{M}} \in \mathscr{M}, \ \mathbf{x}_{\mathscr{M}^{\perp}} \in \mathscr{M}^{\perp}.$$

This orthogonal decomposition is denoted by

$$\mathscr{H} = \mathscr{M} \oplus \mathscr{M}^{\perp}.$$

Moreover, there exists a $P_{\mathscr{M}} \in \mathscr{L}(\mathscr{H})$ such that

$$P_{\mathscr{M}}\mathbf{x} = \mathbf{x}_{\mathscr{M}}, \quad \mathbf{x} \in \mathscr{H}. \tag{A.6.7}$$

This operator $P_{\mathscr{M}}$ is called the *orthogonal projection* (projection for short) of $\mathscr{H}$ onto $\mathscr{M}$. One can show that $P_{\mathscr{M}}$ is a self-adjoint contraction and that $P_{\mathscr{M}}^2 = P_{\mathscr{M}}$. Furthermore,

$$I - P_{\mathscr{M}} = P_{\mathscr{M}^{\perp}}.$$

For example, the Riesz projection $P : L^2 \to L^2$ defined in (A.4.2) is an orthogonal projection.

If $\{\mathbf{x}_j : j \geqslant 1\}$ is an orthonormal basis for $\mathscr{M}$, one can show that

$$P_{\mathscr{M}}\mathbf{x} = \sum_{j \geqslant 1} \langle \mathbf{x}, \mathbf{x}_j \rangle \mathbf{x}_j. \tag{A.6.8}$$

We can also define an orthogonal projection on $\mathscr{H}$ without reference to a subspace. We say a $P \in \mathscr{L}(\mathscr{H})$ is an *orthogonal projection* if $P^2 = P$ and $\ker P = (\operatorname{ran} P)^{\perp}$. In this case, P is the orthogonal projection of $\mathscr{H}$ onto $\operatorname{ran} P$. One can show that if $P \in \mathscr{L}(\mathscr{H})$ satisfies $P^2 = P$, then the following are equivalent (i) P is an orthogonal projection; (ii) $\|P\| = 1$; (iii) $P = P^*$.

Rank-One Operators

For $\mathbf{x}, \mathbf{y} \in \mathscr{H}$, define $\mathbf{x} \otimes \mathbf{y} \in \mathscr{L}(\mathscr{H})$ by

$$(\mathbf{x} \otimes \mathbf{y})(\mathbf{z}) = \langle \mathbf{z}, \mathbf{y} \rangle \mathbf{x}, \quad \mathbf{z} \in \mathscr{H}. \tag{A.6.9}$$

If $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$, then $\mathbf{x} \otimes \mathbf{y}$ is a rank-one operator whose range is span$\{\mathbf{x}\}$. A computation confirms that

$$(\mathbf{x} \otimes \mathbf{y})^* = \mathbf{y} \otimes \mathbf{x}. \tag{A.6.10}$$

If $\|\mathbf{x}\| = 1$, then $\mathbf{x} \otimes \mathbf{x}$ is the orthogonal projection from $\mathscr{H}$ onto span$\{\mathbf{x}\}$.

Spectrum

For $T \in \mathscr{L}(\mathscr{H})$, the *spectrum* $\sigma(T)$ of T is

$$\sigma(T) := \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not invertible in } \mathscr{L}(\mathscr{H})\}.$$

The spectrum is a nonempty compact subset of $\mathbb{C}$ and

$$\sigma(T) \subseteq \{z : |z| \leqslant \|T\|\}.$$

There is also the following finer relationship between $\sigma(T)$ and $\|T\|$.

Theorem A.6.11 (Spectral Radius Formula) *If $T \in \mathscr{L}(\mathscr{H})$, then*

$$\lim_{n\to\infty} \|T^n\|^{1/n} = \sup\{|z| : z \in \sigma(T)\}.$$

Let

$$\sigma_p(T) := \Big\{\lambda \in \mathbb{C} : \ker(\lambda I - T) \neq \{\mathbf{0}\}\Big\}$$

denote the *point spectrum* of T (the eigenvalues of T) and note that $\sigma_p(T) \subseteq \sigma(T)$. Though the spectrum $\sigma(T)$ is always nonempty, the point spectrum $\sigma_p(T)$ might be empty. The *approximate point spectrum* of T is

$$\sigma_{ap}(T) := \{\lambda \in \mathbb{C} : \exists \mathbf{x}_n, \|\mathbf{x}_n\| = 1, \|(\lambda I - T)\mathbf{x}_n\| \to 0\}.$$

One can show that

$$\partial\sigma(T) \subseteq \sigma_{ap}(T) \tag{A.6.12}$$

and hence the approximate point spectrum of an operator is nonempty.

A.7 Toeplitz Operators

For a Lebesgue measurable subset $E \subseteq \mathbb{T}$, let $m(E)$ denote normalized Lebesgue measure of E. By normalized we mean that $m(\mathbb{T}) = 1$. Let L^∞ denote the set of complex-valued Lebesgue measurable functions f on $\mathbb{T}$ whose *essential supremum norm*

$$\|f\|_\infty := \sup \Big\{ a \geqslant 0 : m(\{\xi \in \mathbb{T} : |f(\xi)| > a\}) > 0 \Big\}$$

is finite. Note that

$$\|fg\|_{L^2} \leqslant \|f\|_\infty \|g\|_{L^2}, \quad f \in L^\infty, \quad g \in L^2. \tag{A.7.1}$$

This shows that for each $\phi \in L^\infty$, the multiplication operator

$$M_\phi : L^2 \to L^2, \quad M_\phi f = \phi f,$$

is bounded. Furthermore, one can show that

$$\|M_\phi\| = \|\phi\|_\infty.$$

For $\phi \in L^\infty$, the *Toeplitz operator* $T_\phi \in \mathscr{L}(H^2)$ with *symbol* ϕ is

$$T_\phi := PM_\phi|_{H^2},$$

where P is the Riesz projection from (A.4.2). The (m, n) entry of the matrix representation of T_ϕ with respect to the orthonormal basis $\{1, z, z^2, z^3, \ldots\}$ for H^2 is

$$\langle T_\phi \zeta^m, \zeta^n \rangle = \widehat{\phi}(m - n). \tag{A.7.2}$$

With $\alpha_k = \widehat{\phi}(k)$, this produces an infinite *Toeplitz matrix*

$$\begin{bmatrix} \alpha_0 & \alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \alpha_{-4} & \cdots \\ \alpha_1 & \alpha_0 & \alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \cdots \\ \alpha_2 & \alpha_1 & \alpha_0 & \alpha_{-1} & \alpha_{-2} & \cdots \\ \alpha_3 & \alpha_2 & \alpha_1 & \alpha_0 & \alpha_{-1} & \cdots \\ \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \alpha_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

There is a simple expression for the operator norm of T_ϕ.

Theorem A.7.3 (Brown–Halmos [13]) $\|T_\phi\| = \|\phi\|_\infty$.

Observe that the inequality $\|T_\phi\| \leqslant \|\phi\|_\infty$ follows from (A.7.1). The reverse inequality is more involved.

Riesz Functional Calculus

For $T \in \mathscr{L}(\mathscr{H})$ and a polynomial

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n,$$

define $p(T)$ by

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_n T^n. \tag{A.7.4}$$

Defining $p(T)$ for a wider class of functions p is more subtle. For example, there are convergence issues if the finite series above is replaced by an infinite one. If f is analytic in a neighborhood of $\sigma(T)$ and Γ is a rectifiable Jordan curve (or positively oriented system of Jordan curves) with $\sigma(T)$ inside Γ (positive winding number), the Cauchy Integral formula says

$$f(z) = \frac{1}{n(\Gamma, z)2\pi i} \int_\Gamma \frac{f(\xi)}{\xi - z} d\xi$$

for z inside Γ, where $n(\Gamma, z)$ is the winding number of Γ about z. It makes sense to define $f(T)$ by

$$f(T) := \frac{1}{n(\Gamma, z)2\pi i} \int_\Gamma f(\xi)(\xi I - T)^{-1} d\xi. \tag{A.7.5}$$

The preceding expression is well defined since $\xi \in \mathbb{C} \backslash \sigma(T)$. This definition of $f(T)$ agrees with the definition of $f(T)$ from (A.7.4) when f is a polynomial. There is also the following relationship between $\sigma(f(T))$ and $f(\sigma(T))$.

Theorem A.7.6 (Spectral Mapping Theorem) *If f is analytic in a neighborhood of $\sigma(T)$, then $\sigma(f(T)) = f(\sigma(T))$.*

This definition of $f(T)$ for any $T \in \mathscr{L}(\mathscr{H})$ and f analytic in a neighborhood of $\sigma(T)$ is the *Riesz functional calculus*. If one places further restrictions on T (normal, contraction, etc.), one can sometimes extend the class of functions f for which one can meaningfully define $f(T)$.

Unitary Equivalence

If $\mathscr{H}_1$ and $\mathscr{H}_2$ are Hilbert spaces, we say that $T_1 \in \mathscr{L}(\mathscr{H}_1)$ and $T_2 \in \mathscr{H}_2$ are *unitarily equivalent* if there is a linear isometry U from $\mathscr{H}_1$ onto $\mathscr{H}_2$ such that

$$T_2 U = U T_1. \tag{A.7.7}$$

Note that unitarily equivalent operators have the same norm, eigenvalues, and spectrum.

A.8 Schur's Theorem

If $A \in M_n$ is a self-adjoint matrix, that is $A = A^*$, the spectral theorem says that A is unitarily equivalent to the diagonal matrix

$$\begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_{n-1} & \\ & & & & \lambda_n \end{bmatrix},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A. For a general matrix $A \in M_n$, we have the following [79].

Theorem A.8.1 (Schur's Theorem) *If $A \in M_n$ has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, in any prescribed order, then A is unitarily equivalent to an upper-triangular matrix with main diagonal $\lambda_1, \lambda_2, \ldots, \lambda_n$.*

References

1. J. Agler, J.E. McCarthy, *Pick Interpolation and Hilbert Function Spaces*. Graduate Studies in Mathematics, vol. 44 (American Mathematical Society, Providence, 2002)
2. L.V. Ahlfors, An extension of Schwarz's lemma. Trans. Am. Math. Soc. **43**(3), 359–364 (1938)
3. D.Z. Arov, Darlington's method in the study of dissipative systems. Dokl. Akad. Nauk SSSR **201**(3), 559–562 (1971)
4. F. Bagemihl, G. Piranian, Boundary functions for functions defined in a disk. Mich. Math. J. **8**, 201–207 (1961)
5. A.F. Beardon, D. Minda, A multi-point Schwarz-Pick lemma. J. Anal. Math. **92**, 81–104 (2004)
6. A.F. Beardon, T.K. Carne, T.W. Ng, The critical values of a polynomial. Constr. Approx. **18**(3), 343–354 (2002)
7. C.A. Berger, J.G. Stampfli, Mapping theorems for the numerical range. Am. J. Math. **89**, 1047–1055 (1967)
8. S.J. Bernau, F. Smithies, A note on normal operators. Proc. Camb. Philos. Soc. **59**, 727–729 (1963)
9. H. Bohr, A theorem concerning power series. Proc. Lond. Math. Soc. **S2-13**(1), 1–5 (1914)
10. E. Bombieri, Sopra un teorema di H. Bohr e G. Ricci sulle funzioni maggioranti delle serie di potenze. Boll. Un. Mat. Ital. (3) **17**, 276–282 (1962)
11. E. Bombieri, J. Bourgain, A remark on Bohr's inequality. Int. Math. Res. Not. (80), 4307–4330 (2004)
12. L.E.J. Brouwer, Zur Invarianz des n-dimensionalen Gebiets. Math. Ann. **72**(1), 55–56 (1912)
13. A. Brown, P.R. Halmos, Algebraic properties of Toeplitz operators. J. Reine Angew. Math. **213**, 89–102 (1963/1964)
14. D.G. Cantor, R.R. Phelps, An elementary interpolation theorem. Proc. Am. Math. Soc. **16**, 523–525 (1965)
15. C. Carathéodory, Untersuchungen über die konformen Abbildungen von festen und veränderlichen Gebieten. Math. Ann. **72**(1), 107–144 (1912)
16. C. Carathéodory, *Theory of Functions of a Complex Variable*, vol. 2 (Chelsea Publishing Company, New York, 1954). Translated by F. Steinhardt
17. C. Carathéodory, H. Rademacher, Über die Eineindentigkeit im Kleinen und im Grossen Stetiger Abdildungen von Gebieten. Arch. Math. Phys. **26**(3), 1–9 (1917)
18. G. Cassier, I. Chalendar, The group of the invariants of a finite Blaschke product. Complex Var. Theory Appl. **42**(3), 193–206 (2000)

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8

19. I. Chalendar, R. Mortini, When do finite Blaschke products commute? Bull. Aust. Math. Soc. **64**(2), 189–200 (2001)
20. D. Challener, L. Rubel, A converse to Rouché's theorem. Am. Math. Mon. **89**(5), 302–305 (1982)
21. I. Chalendar, P. Gorkin, J.R. Partington, Numerical ranges of restricted shifts and unitary dilations. Oper. Matrices **3**(2), 271–281 (2009)
22. I. Chalendar, P. Gorkin, J.R. Partington, Prime and semiprime inner functions. J. Lond. Math. Soc. (2) **88**(3), 779–800 (2013)
23. I. Chalendar, P. Gorkin, J. Partington, W. T. Ross, Clark measures and a theorem of Ritt. Math. Scand. **122**, 277–298 (2018)
24. D.N. Clark, One dimensional perturbations of restricted shifts. J. Anal. Math. **25**, 169–191 (1972)
25. E.F. Collingwood, A.J. Lohwater, *The Theory of Cluster Sets*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 56 (Cambridge University Press, Cambridge, 1966)
26. J.B. Conway, *Functions of One Complex Variable*. Graduate Texts in Mathematics, vol. 11, 2nd edn. (Springer, New York, 1978)
27. J.B. Conway, *A Course in Functional Analysis*. Graduate Texts in Mathematics, vol. 96, 2nd edn. (Springer, New York, 1990)
28. C. Cowen, Finite Blaschke products and compositions of other finite Blaschke products. arXiv:1207.4010
29. U. Daepp, P. Gorkin, R. Mortini, Ellipses and finite Blaschke products. Am. Math. Mon. **109**(9), 785–795 (2002)
30. U. Daepp, P. Gorkin, A. Shaffer, B. Sokolowsky, K. Voss, Decomposing finite Blaschke products. J. Math. Anal. Appl. **426**(2), 1201–1216 (2015)
31. S. Dineen, *The Schwarz Lemma*. Oxford Mathematical Monographs (The Clarendon Press Oxford University Press/Oxford Science Publications, New York, 1989)
32. P.G. Dixon, Banach algebras satisfying the non-unital von Neumann inequality. Bull. Lond. Math. Soc. **27**(4), 359–362 (1995)
33. R.G. Douglas, J. William Helton, Inner dilations of analytic matrix functions and Darlington synthesis. Acta Sci. Math. (Szeged) **34**, 61–67 (1973)
34. R.G. Douglas, H.S. Shapiro, A.L. Shields, Cyclic vectors and invariant subspaces for the backward shift operator. Ann. Inst. Fourier (Grenoble) **20**(fasc. 1), 37–76 (1970)
35. M.A. Dritschel, J. Rovnyak, The operator Fejér-Riesz theorem, in *A Glimpse at Hilbert Space Operators*. Operator Theory: Advances and Applications, vol. 207 (Birkhäuser Verlag, Basel, 2010), pp. 223–254
36. S.W. Drury, Symbolic calculus of operators with unit numerical radius. Linear Algebra Appl. **428**(8–9), 2061–2069 (2008)
37. P. du Bois-Reymond, Ueber die Fourierschen Reihen. Nachr. Kön. Ges. Wiss. Göttingen **21**, 571–582 (1873)
38. P.L. Duren, *Theory of H^p Spaces* (Academic, New York, 1970)
39. P.L. Duren, Univalent Functions. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 259 (Springer, New York, 1983)
40. E. Egerváry, On the contractive linear transformations of n-dimensional vector space. Acta Sci. Math. Szeged **15**, 178–182 (1954)
41. I. Erdélyi, On partial isometries in finite-dimensional Euclidean spaces. SIAM J. Appl. Math. **14**, 453–467 (1966)
42. Euclid, *Euclid's Elements* (Green Lion Press, Santa Fe, 2002). All Thirteen Books Complete in One Volume, The Thomas L. Heath translation, Edited by Dana Densmore
43. P. Fatou, Sur les équations fonctionnelles. Bull. Soc. Math. Fr. **47**, 161–271 (1919)
44. P. Fatou, Sur les équations fonctionnelles. Bull. Soc. Math. Fr. **48**, 208–314 (1920)
45. P. Fatou, Sur les équations fonctionnelles. Bull. Soc. Math. Fr. **48**, 33–94 (1920)
46. P. Fatou, Sur les fonctions holomorphes et bornées à l'intérieur d'un cercle. Bull. Soc. Math. Fr. **51**, 191–202 (1923)

47. L. Fejér, Über trigonometrische Polynome. J. Reine Angew. Math. **146**, 53–82 (1916)
48. S. Fisher, The convex hull of the finite Blaschke products. Bull. Am. Math. Soc. **74**, 1128–1129 (1968)
49. E. Fricain, J. Mashreghi, On a characterization of finite Blaschke products. Complex Var. Elliptic Equ. **59**(3), 362–368 (2014)
50. O. Frostman, Sur les produits de Blaschke. Kungl. Fysiografiska Sällskapets i Lund Förhandlingar [Proc. R. Physiog. Soc. Lund] **12**(15), 169–182 (1942)
51. S.R. Garcia, Conjugation, the backward shift, and Toeplitz kernels. J. Oper. Theory **54**(2), 239–250 (2005)
52. S.R. Garcia, Inner matrices and Darlington synthesis. Methods Funct. Anal. Topol. **11**(1), 37–47 (2005)
53. S.R. Garcia, Conjugation and Clark operators, in *Recent Advances in Operator-Related Function Theory*. Contemporary Mathematics, vol. 393 (American Mathematical Society, Providence, 2006), pp. 67–111
54. S.R. Garcia, M. Putinar, Complex symmetric operators and applications. Trans. Am. Math. Soc. **358**(3), 1285–1315 (electronic) (2006)
55. S.R. Garcia, M. Putinar, Complex symmetric operators and applications. II. Trans. Am. Math. Soc. **359**(8), 3913–3931 (electronic) (2007)
56. S.R. Garcia, D. Sarason, Real outer functions. Indiana Univ. Math. J. **52**(6), 1397–1412 (2003)
57. S.R. Garcia, D. Sherman, Matrices similar to partial isometries. Linear Algebra Appl. **526**, 35–41 (2017)
58. S.R. Garcia, E. Prodan, M. Putinar, Mathematical and physical aspects of complex symmetric operators. J. Phys. A **47**(35), 353001, 54 (2014)
59. S.R. Garcia, J. Mashreghi, W.T. Ross, *Introduction to Model Spaces and Their Operators*. Cambridge Studies in Advanced Mathematics, vol. 148 (Cambridge University Press, Cambridge, 2016)
60. S.R. Garcia, J. Mashreghi, W.T. Ross, Real complex functions, in *Recent Progress on Operator Theory and Approximation in Spaces of Analytic Functions*. Contemporary Mathematics, vol. 679 (American Mathematical Society, Providence, 2016), pp. 91–128
61. J. Garnett, *Bounded Analytic Functions*. Graduate Texts in Mathematics, vol. 236, 1st edn. (Springer, New York, 2007)
62. H.-L. Gau, P.Y. Wu, Numerical range of $S(\phi)$. Linear Multilinear Algebra **45**(1), 49–73 (1998)
63. H.-L. Gau, P.Y. Wu, Finite Blaschke products of contractions. Linear Algebra Appl. **368**, 359–370 (2003)
64. I. Glicksberg, A remark on Rouché's theorem. Am. Math. Mon. **83**(3), 186–187 (1976)
65. P. Gorkin, R.C. Rhoades, Boundary interpolation by finite Blaschke products. Constr. Approx. **27**(1), 75–98 (2008)
66. P. Gorkin, L. Laroco, R. Mortini, R. Rupp, Composition of inner functions. Results Math. **25**(3–4), 252–269 (1994)
67. P.R. Halmos, J.E. McLaughlin, Partial isometries. Pacific J. Math. **13**, 585–596 (1963)
68. R. Hartshorne, *Geometry: Euclid and Beyond*. Undergraduate Texts in Mathematics (Springer, New York, 2000)
69. F. Hausdorff, Der Wertvorrat einer Bilinearform. Math. Z. **3**(1), 314–316 (1919)
70. W.K. Hayman, Coefficient problems for univalent functions and related function classes. J. Lond. Math. Soc. **40**, 385–406 (1965)
71. J.Z. Hearon, Partially isometric matrices. J. Res. Nat. Bur. Stand. Sect. B **71B**, 225–228 (1967)
72. M. Heins, A residue theorem for finite Blaschke products. Proc. Am. Math. Soc. **2**, 622–624 (1951)
73. M. Heins, *Selected Topics in the Classical Theory of Functions of a Complex Variable*. Athena Series: Selected Topics in Mathematics (Holt, Rinehart and Winston, New York, 1962)

74. M. Heins, Some characterizations of finite Blaschke products of positive degree. J. Anal. Math. **46**, 162–166 (1986)
75. H. Helson, Large analytic functions, in *Linear Operators in Function Spaces (Timişoara, 1988)*. Operator Theory: Advances and Applications, vol. 43 (Birkhäuser, Basel, 1990), pp. 209–216
76. H. Helson, Large analytic functions. II, in *Analysis and Partial Differential Equations*. Lecture Notes in Pure and Applied Mathematics, vol. 122 (Dekker, New York, 1990), pp. 217–220
77. H. Helson, D. Sarason, Past and future. Math. Scand **21**, 5–16 (1967/1968)
78. G.A. Hjelle, Constructing interpolating Blaschke products with given preimages. Comput. Methods Funct. Theory **7**(1), 43–54 (2007)
79. R.A. Horn, C.R. Johnson, *Matrix Analysis*, 2nd edn. (Cambridge University Press, Cambridge, 2013)
80. A.L. Horwitz, L.A. Rubel, A uniqueness theorem for monic Blaschke products. Proc. Am. Math. Soc. **96**(1), 180–182 (1986)
81. N. Jacobson, *Basic Algebra. I*, 2nd edn. (W.H. Freeman and Company, New York, 1985)
82. W.B. Jones, S. Ruscheweyh, Blaschke product interpolation and its application to the design of digital filters. Constr. Approx. **3**(4), 405–409 (1987)
83. G. Julia, Principes géométriques d'analyse, in *Première partie* (Gauthier-Villars, Paris, 1930)
84. T.J. Kaczynski, Boundary functions for function defined in a disk. J. Math. Mech. **14**, 589–612 (1965)
85. T.J. Kaczynski, On a boundary property of continuous functions. Mich. Math. J. **13**, 313–320 (1966)
86. D. Kalman, An elementary proof of Marden's theorem. Am. Math. Mon. **115**(4), 330–338 (2008)
87. D. Kincaid, W. Cheney, *Numerical Analysis*, 2nd edn. (Brooks/Cole Publishing Co./Mathematics of Scientific Computing, Pacific Grove, 1996)
88. H. Klaja, J. Mashreghi, T. Ransford, On mapping theorems for numerical range. Proc. Am. Math. Soc. **144**(7), 3009–3018 (2016)
89. P. Koebe, Über die Uniformisierung der algebraischen Kurven. II. Math. Ann. **69**(1), 1–81 (1910)
90. D. Kraus, O. Roth, Critical points, the Gauss curvature equation and Blaschke products, in *Blaschke Products and Their Applications*. Fields Institute Communications, vol. 65 (Springer, New York, 2013), pp. 133–157
91. M.G. Krein, M.A. Naimark, The method of symmetric and Hermitian forms in the theory of the separation of the roots of algebraic equations. Linear Multilinear Algebra **10**(4), 265–308 (1981). Translated from the Russian by O. Boshko and J. L. Howland
92. E. Landau, *Vorlesungen uber Zahlentheorie*, 1st edn. (Verlag Von S. Hirzel, 1927)
93. P.D. Lax, B. Wendroff, Systems of conservation laws. Commun. Pure Appl. Math. **13**, 217–237 (1960)
94. P.D. Lax, B. Wendroff, On the stability of difference schemes. Commun. Pure Appl. Math. **15**, 363–371 (1962)
95. E. Levy, O.M. Shalit, Dilation theory in finite dimensions: the possible, the impossible and the unknown. Rocky Mt. J. Math. **44**(1), 203–221 (2014)
96. C.-K. Li, A simple proof of the elliptical range theorem. Proc. Am. Math. Soc. **124**(7), 1985–1986 (1996)
97. M.S. Livšic, Isometric operators with equal deficiency indices, quasi-unitary operators. Am. Math. Soc. Transl. (2) **13**, 85–103 (1960)
98. S. Łojasiewicz, *Introduction to Complex Analytic Geometry* (Birkhäuser Verlag, Basel, 1991). Translated from the Polish by Maciej Klimek.
99. M. Marden, *Geometry of Polynomials*, 2nd edn. Mathematical Surveys, No. 3 (American Mathematical Society, Providence, 1966)
100. A.I. Markushevich, *Theory of Functions of a Complex Variable*, vols. I–III, English edn. (Chelsea Publishing Co., New York, 1977). Translated and edited by Richard A. Silverman

101. J. Mashreghi, *Representation Theorems in Hardy Spaces*. London Mathematical Society Student Texts, vol. 74 (Cambridge University Press, Cambridge, 2009)
102. J.E. McMillan, Boundary properties of functions continuous in a disc. Mich. Math. J. **13**, 299–312 (1966)
103. D. Minda, G. Schober, Another elementary approach to the theorems of Landau, Montel, Picard and Schottky. Complex Var. Theory Appl. **2**(2), 157–164 (1983)
104. Y.A. Muhanna, R.M. Ali, S. Ponnusamy, On the Bohr inequality. arXiv:1612.00597
105. J.R. Munkres, *Elements of Algebraic Topology* (Addison-Wesley Publishing Company, Menlo Park, 1984)
106. N. Nikolski, *Treatise on the Shift Operator* (Springer, Berlin, 1986)
107. D.J. Newman, H.S. Shapiro, The Taylor coefficients of inner functions. Mich. Math. J. **9**, 249–255 (1962)
108. T.W. Ng, C.Y. Tsang, Polynomials versus finite Blaschke products, in *Blaschke Products and Their Applications*. Fields Institute Communications, vol. 65 (Springer, New York, 2013), pp. 249–273
109. T.W. Ng, M.-X. Wang, Ritt's theory on the unit disk. Forum Math. **25**(4), 821–851 (2013)
110. V.I. Paulsen, G. Popescu, D. Singh, On Bohr's inequality. Proc. Lond. Math. Soc. (3) **85**(2), 493–512 (2002)
111. V.I. Paulsen, M. Raghupathi, *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*. Cambridge Studies in Advanced Mathematics, vol. 152 (Cambridge University Press, Cambridge, 2016)
112. C. Pearcy, An elementary proof of the power inequality for the numerical radius. Mich. Math. J. **13**, 289–291 (1966)
113. A.A. Pekarskiĭ, Estimates of the derivative of a Cauchy-type integral with meromorphic density and their applications. Mat. Zametki **31**(3), 389–402, 474 (1982)
114. G. Pick, Über die Beschränkungen analytischer Funktionen, welche durch vorgegebene Funktionswerte bewirkt werden. Math. Ann. **77**(1), 7–23 (1915)
115. G. Pisier, *Similarity Problems and Completely Bounded Maps*. Lecture Notes in Mathematics, vol. 1618, expanded edn. (Springer, Berlin, 2001). Includes the Solution to "The Halmos Problem"
116. T. Radó, Zur theorie der mehrdeutigen konformen abbildungen. Acta Litt. Scient. Univ. Hung. (Szeged) (1), 55–64 (1922–1923)
117. G. Ricci, Complementi a un teorema di H. Bohr riguardante le serie di potenze. Rev. Un. Mat. Argentina **17**, 185–195 (1955/1956)
118. F. Riesz, über ein Probelm des Herrn Carathéodory. J. Reine Angew. Math. **146**, 83–87 (1916)
119. J.F. Ritt, Prime and composite polynomials. Trans. Am. Math. Soc. **23**(1), 51–66 (1922)
120. P. Rizzonelli, Valutazioni del tipo di H. Bohr per le maggioranti delle serie de potenze. Riv. Mat. Univ. Parma (2) **3**, 259–270 (1962)
121. W.T. Ross, H.S. Shapiro, *Generalized Analytic Continuation*. University Lecture Series, vol. 25 (American Mathematical Society, Providence, 2002)
122. H.L. Royden, *Real Analysis*, 3rd edn. (Macmillan Publishing Company, New York, 1988)
123. W. Rudin, *Real and Complex Analysis* (McGraw-Hill Book Co., New York, 1966)
124. G. Schmeisser, The conjectures of Sendov and Smale, in *Approximation Theory* (DARBA, Sofia, 2002), pp. 353–369
125. H.A. Schwarz, *Gesammelte mathematische Abhandlungen. Band I, II* (Chelsea Publishing Co., Bronx, 1972). Nachdruck in einem Band der Auflage von 1890
126. S. Sidon, Über einen Satz von Herrn Bohr. Math. Z. **26**(1), 731–732 (1927)
127. W. Sierpinski, Sur une série potentielle qui, étant convergente en tout point son cercle convergence, représente sur ce cercle une fonction discontinue. Rend. Circ. Mat. Palermo **41**, 187–190 (1916)
128. B. Simon, *Advanced Complex Analysis*. A Comprehensive Course in Analysis, Part 2B (American Mathematical Society, Providence, 2015)
129. D.A. Singer, The location of critical points of finite Blaschke products. Conform. Geom. Dyn. **10**, 117–124 (electronic) (2006)

130. B. Sokolowsky, Decomposing Blaschke products and polynomials. Master's Thesis. Bucknell University, Lewisburg (May 2013)
131. B. Szőkefalvi-Nagy, C. Foias, H. Bercovici, L. Kérchy, *Harmonic Analysis of Operators on Hilbert Space*, 2nd edn. Universitext (Springer, New York, 2010)
132. S. Takenaka, On the orthonormal functions and a new formula of interpolation. Jap. J. Math. **2**, 129–145 (1925)
133. O. Toeplitz, Das algebraische Analogon zu einem Satze von Fejér. Math. Z. **2**(1–2), 187–197 (1918)
134. M. Tomić, Sur un théorème de H. Bohr. Math. Scand. **11**, 103–106 (1962)
135. B.L. van der Waerden, *Modern Algebra*, vol. I (Frederick Ungar Publishing Co., New York, 1949). Translated from the second revised German edition by Fred Blum, With revisions and additions by the author
136. J. von Neumann, Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes. Math. Nachr. **4**, 258–281 (1951)
137. J.L. Walsh, Note on the location of zeros of extremal polynomials in the non-euclidean plane. Acad. Serbe Sci. Publ. Inst. Math. **4**, 157–160 (1952)
138. R. Younis, Interpolation by a finite Blaschke product. Proc. Am. Math. Soc. **78**(3), 451–452 (1980)
139. S. Zakeri, On critical points of proper holomorphic maps on the unit disk. Bull. Lond. Math. Soc. **30**(1), 62–66 (1998)

Index

S. R. Garcia et al., *Finite Blaschke Products and Their Connections*,
https://doi.org/10.1007/978-3-319-78247-8