Progress in Mathematics
341

John P. D'Angelo

Rational Sphere Maps

Birkhäuser

Birkhäuser

Progress in Mathematics

Volume 341

More information about this series at http://www.springer.com/series/4848

John P. D'Angelo

Rational Sphere Maps

Birkhäuser

John P. D'Angelo
Department of Mathematics
University of Illinois, Urbana-Champaign
Urbana, IL, USA

ISSN 0743-1643 ISSN 2296-505X (electronic)
Progress in Mathematics
ISBN 978-3-030-75808-0 ISBN 978-3-030-75809-7 (eBook)
https://doi.org/10.1007/978-3-030-75809-7

Mathematics Subject Classification: 32M99, 32A10, 15B57, 32V99

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The unit circle S^1 in the complex number system $\mathbb{C}$ and its self-mappings have played a major role in the history of mathematics. Below we give many striking examples. The central theme throughout this book will be to understand higher dimensional analogues, where things are more subtle and ideas from many fields of mathematics make their appearance.

In one dimension, if f is holomorphic (complex analytic) in a neighborhood of the closure of the unit disk $\mathbb{B}_1$, and f maps the circle to itself, then f is a finite Blaschke product. One can draw the same conclusion assuming only that f is a proper holomorphic mapping from $\mathbb{B}_1$ to itself. In particular such functions are rational. Our primary topic will be the study of holomorphic rational maps sending the unit sphere in the source complex Euclidean space $\mathbb{C}^n$ to the unit sphere in some target space $\mathbb{C}^N$. We call such mappings *rational sphere maps*. We use the terms monomial sphere map and polynomial sphere map with obvious meaning; even these mappings exhibit remarkably interesting and complicated behavior as the source and target dimensions rise.

In this book, a *rational sphere map* f is complex analytic where it is defined. In other words, f depends on the z variables but not on the $\bar{z}$ variables. In Chap. 6 we briefly discuss some differences between holomorphic polynomial sphere maps and real polynomial sphere maps. In particular, in complex dimension n at least 2, the only non-constant holomorphic polynomial maps sending the unit sphere to itself are linear, whereas there are real polynomial sphere maps of every degree. I considered the title *Complex Analytic Rational Sphere Maps* to prevent possible confusion, but the shorter title seems more appealing.

In some sense, this book is a research monograph, as it develops in a systematic fashion most of the research on rational sphere maps done in the last forty years. It differs however from many monographs in several ways, which we now describe.

First of all, scattered throughout the book are a large number of computational examples; the author feels that merging the abstract and concrete enhances both. Many times in his work on this subject, a theorem resulted from trying to cast a collection of examples into one framework. Some readers will stare at these formulas, observe subtle patterns, and pose their own open questions. Other readers may find the formulas distracting. I hope that I have achieved the right balance. Chaps. 3 and 4 include formulas that could not easily be obtained by hand computation. Mathematica was used to help perform some of these calculations. The author acknowledges assistance in coding received from Jiri Lebl, Daniel Lichtblau, Dan Putnam, and Bob Vanderbei. Some results from coding have led to theorems and others have led to unanswered questions. Both types of results appear here. Section 4.9 includes recent code by Lichtblau [1].

Second, I have included more than 100 exercises. Most of these are computational and have a simple purpose: give the reader something to do when things become confusing. These exercises are numbered by Chapter and often appear in the middle of a section. Given the many search tools available, this method seems most appropriate. This book hopes to expose some beautiful mathematics; it is not a calculus text where long lists of exercises appear at the end of each section. The exercises are meant for readers who enjoy them but none are indispensable to the general development.

I have posed fifteen open problems here. They belong to many parts of mathematics; the symmetry of the unit sphere is responsible for their variety. These problems appear within the text but are repeated in a short chapter at the end of the book. The author hopes that this book will enhance research by engaging others in both what is known and where this knowledge leads.

Section 1.7 provides a kind of global positioning system for the book. It locates where in the book some of the fundamental results are discussed and indicates what happens in each chapter. The author modestly hopes that both experts and novices find this *map* to be useful both in learning about rational sphere maps and navigating the book.

To introduce the subject of rational sphere maps, we provide several examples in one dimension and indicate how to extend the ideas to higher dimensions.

Example 1 Many elementary trigonometric identities are easily proved by combining the binomial expansion with de Moivre's formula

$$(\cos(\theta) + i\sin(\theta))^m = (e^{i\theta})^m = e^{im\theta} = \cos(m\theta) + i\sin(m\theta). \qquad (*)$$

In fact every trig identity follows from the following facts:

1. The complex numbers are a field.
2. $\cos(z) = \frac{e^{iz}+e^{-iz}}{2}$ and $\sin(z) = \frac{e^{iz}+e^{-iz}}{2i}$.
3. For complex numbers z, w we have $e^{z+w} = e^z e^w$.
4. Complex conjugation is continuous, and hence $e^{\bar{z}} = \bar{e^z}$.

It is natural to take (2) as the definition of the trig functions. Combining (2) and (3) yields $\cos^2(z) + \sin^2(z) = 1$. Combining (3) and (4) yields $|e^{i\theta}|^2 = 1$ when θ is real. Item (4) is needed because the exponential function is defined by its power series; one needs to know that the conjugate of a convergent infinite sum is the infinite sum of the conjugated terms.

Formula (*) is closely related to the map $z \to z^m$, which sends the circle to itself, and hence is a monomial sphere map. One higher dimensional analogue of this mapping will be the tensor product $z \to z^{\otimes m}$ for $z \in \mathbb{C}^n$. The tensor product provides a monomial sphere map, but requires a higher dimensional target space. We will encounter *restricted tensor products* and a kind of *tensor division*.

Example 2 The unit circle can be regarded as the unitary group $\mathrm{U}(1)$. The m-th roots of unity form a finite cyclic subgroup Γ_m under multiplication. The map $z \to z^m$ sends the circle to itself and is invariant under Γ_m. We will study analogues in higher dimensions in Chap. 5, by associating both invariant and equivariant groups with rational sphere maps. The unitary group $\mathrm{U}(n)$ and the holomorphic automorphism group of the unit ball arise throughout. In addition, representations of Γ_m in $\mathrm{U}(2)$ lead in Chap. 3 to interesting combinatorial results.

Example 3 The theory of Fourier series is based upon the complete orthonormal system $\{e^{im\theta}\}$ for $L^2(S^1)$. Closely related is the result that the monomials $z \to z^m$ form a complete orthogonal system for $L^2(\mathbb{B}_1)$. The analogous statement for the monomials $z \to z^\alpha$ holds in any dimension.

Example 4 Riemann surfaces arose from trying to visualize the space of solutions to equations such as $z^m = w$. We will study proper mappings from $\mathbb{B}_n$ to $\mathbb{B}_N$; the image of the ball is then an n-dimensional complex variety. We also study a subvariety of $\mathbb{B}_n \times \mathbb{C}^N$ associated with a rational sphere map. This variety contains the graph of the map, but exceptional fibers often arise.

Example 5 Each factor (including the $e^{i\theta}$ term) of the Blaschke product

$$e^{i\theta} \prod_{j=1}^{m} \frac{a_j - z}{1 - \bar{a}_j z}$$

can be regarded as an automorphism of the unit disk. In n dimensions, the automorphism group of the unit ball $\mathbb{B}_n$ is the Lie group $\mathbf{SU}(n, 1)$ divided by its center. We will see tensor products of automorphisms, but (as in Example 1) new phenomena arise. Not every rational sphere map is a tensor product of automorphisms.

Example 6 Example 5 shows that every polynomial q that does not vanish on the closed unit disk is the denominator of a rational sphere map that is reduced to lowest terms. Proving the analogous statement in higher dimensions is much more subtle and seems to require Hermitian analogues of Hilbert's 17-th problem. Let us elaborate. Suppose $z \in \mathbb{C}^n$ and $r(z, \bar{z})$ is a real-valued polynomial. When the values of r are non-negative, we naturally ask whether r is a Hermitian sum of squares; that is, can we write

$$r(z, \bar{z}) = \sum_{j=1}^{k} |f_j(z)|^2$$

for (holomorphic) polynomials f_j? The answer is *not necessarily*. What can we say? The resulting ideas (see Chap. 2) enable us to prove the following result. Let q

be a polynomial that does not vanish on the closed unit ball in $\mathbb{C}^n$. Then there is an integer N and a polynomial mapping $p : \mathbb{C}^n \to \mathbb{C}^N$ such that $\frac{p}{q}$ is reduced to lowest terms and defines a rational sphere map. There are no bounds possible on N nor on the degree of p that depend only upon n and the degree of q. We emphasize that the easy proof in one dimension does not require these ideas. This discussion combines with Hermitian linear algebra to give Theorem 2.15, which provides a general description of all rational sphere maps.

Example 7 In Chap. 3 we will introduce a class of polynomials in two variables that arise from considering group-invariant monomial sphere maps. These polynomials turn out to be related to Chebyshev polynomials and they exhibit a long list of remarkable properties. One of these properties is that the so-called *freshman's dream*: $(x+y)^d$ is congruent to $x^d + y^d$ modulo d if and only if $d = 1$ or d is prime, holds for these polynomials $f_d(x, y)$ as well.

Example 8 In Chap. 7 we establish a sharp bound on the volume of the image of a polynomial sphere map. A one-dimensional version of this result is quite appealing and we discuss it in detail as well.

The underlying theme in this book derives from the following simple observations. First, the collection of rational sphere maps with a given source dimension n and target dimension N has little algebraic structure, unless $n = N$. In Chap. 2, we show that there is considerably more structure to the problem if we regard the target dimension as a variable. Determining the rational sphere maps of degree d in source dimension n and unspecified target dimension leads to a system of linear equations for the inner products of unknown vectors. See Theorem 2.15. If we assume these vectors are orthogonal, then we obtain a linear system for unknown non-negative numbers. This case is equivalent to the study of monomial sphere maps, which we investigate in Chaps. 3 and 4. Even in the monomial case, the dimension of the set of solutions tends to infinity as the degree tends to infinity.

As usual in Mathematics, when there are too many solutions to a problem, one can restrict the solutions by optimizing various quantities. For example, in Chap. 7, we discuss the volume of the image of the ball under a polynomial sphere map of degree d. We show that the homogeneous mapping $z^{\otimes d}$ provides the *maximum* volume. Chaps. 3 and 4 consider *minimizing* two somewhat related quantities for monomial sphere maps of degree d in source dimension n. One of these quantities is the minimum target dimension; the other is the minimum value of the map at the point with coordinates all equal to 1. We obtain some rather difficult combinatorial and asymptotic results about these problems.

Let us a say a few words about prerequisites. The author believes that everything in this book should be accessible to most mathematicians, including graduate students. Because of the symmetry of the unit sphere, however, the material interacts with nearly all fields of mathematics. We use basic facts from complex analysis, linear algebra, functional analysis, and algebra. We will sometimes use ideas from elementary differential geometry and we will employ combinatorial reasoning. No deep theorems are required. The only prerequisite is appreciation of the ideas.

Numbering in this book is done by Chapter. Thus, for example, Proposition 1.5 means the fifth proposition in Chap. 1. It precedes Corollary 1.1, because there are no items called *corollary* before it. We do not number every displayed equation. This point is worth elaborating. Paul Halmos (I don't know the precise reference, but he said so!) once suggested that every equation should be numbered, because even if the author never refers to a given equation, someone else might. On the other hand, numbering everything seems to clutter things too much. I hope, unrealistically of course, that I have compromised by numbering an equation if and only if it should be numbered.

I wish to acknowledge various mathematicians who have contributed to my understanding of the ideas in this book, or who have coded some computations: Eric Bedford, Dan Burns, Paulo Cordaro, Peter Ebenfelt, Jim Faran, Franc Forstnerič, Dusty Grundmeier, Zhenghui Huo, Bernhard Lamel, Jiri Lebl, Daniel Lichtblau, Han Peters, Dan Putnam, Bob Vanderbei, and Ming Xiao. I also acknowledge Simon Kos, a physicist, who made an important contribution to the ideas in Chap. 3. Quite a few of the results in this book are outgrowths of work I did with Jiri Lebl and other work I did with Ming Xiao. Their contributions have been indispensable. Many other mathematicians have indirectly contributed, primarily via their own inspiring work. I have also benefited from attending meetings and conferences over the years. Let me specifically mention programs at the American Institute of Math, workshops in Serra Negra (Brazil), conferences at the Erwin Schrödinger Institute (Vienna, Austria), and various special sessions at AMS sectional meetings.

I thank Chris Tominich of Springer for his role as editor and especially for his solicitation of useful reviews. I thank the anonymous reviewers for their comments; I modestly hope that I have improved the book by dealing with their suggestions and criticisms.

During the preparation of this book I have been supported by NSF Grants DMS-1066177 and DMS 13-61001. I also acknowledge support from the Kenneth D. Schmidt Professorial Scholar award from the University of Illinois.
I dedicate this book to my wife Annette and our four children John, Lucie, Paul, and Henry.

Urbana, USA John P. D'Angelo

Reference

1. D. Lichtblau, personal communication.

Contents

Chapter 1
Complex Euclidean Space

This chapter develops basic properties of complex Euclidean space. Some of the main ideas are unitary transformations, the holomorphic automorphism group of the unit ball, the use of Hermitian forms, and proper holomorphic mappings. We also gather some elementary combinatorial information.

1 Generalities

The notation $\mathbb{C}^n$ denotes complex Euclidean space of dimension n. As a set, it consists of n-tuples of complex numbers $z = (z_1, ..., z_n)$. The notation includes the information that $\mathbb{C}^n$ is an **inner product space**. The **inner product** of vectors z and w is defined by

$$\langle z, w \rangle = \sum_{j=1}^{n} z_j \overline{w_j}.$$

We denote the corresponding squared norm by

$$\|z\|^2 = \langle z, z \rangle = \sum_{j=1}^{n} |z_j|^2.$$

The set $\mathbb{C}^n$ is then a metric space with distance function given by $\|z - w\|$. As a consequence, we have all the usual notions from point-set topology. In particular, a subset Ω is *open* if, for each $p \in \Omega$, there is a positive ϵ such that $\|z - p\| < \epsilon$ implies $z \in \Omega$. We denote the **unit ball**, centered at the origin, by $\mathbb{B}_n$. Its boundary is the **unit sphere** S^{2n-1}. Odd dimensional spheres arise throughout this book. We assume the reader knows such terms as *connected*, *compact*, *limit*, *sequence*, *subsequence* and

J. P. D'Angelo, *Rational Sphere Maps*, Progress in Mathematics 341,
https://doi.org/10.1007/978-3-030-75809-7_1

so on. A *domain* will be an open, connected set. As a metric space $\mathbb{C}^n$ is complete, and hence $\mathbb{C}^n$ is a *Hilbert space* of dimension n.

Remark 1.1 We use the notations $\|z\|$ and $\langle z, w\rangle$ for the norm and the inner product in any (unspecified) dimension. The implication $\|f(z)\|^2 = 1$ on $\|z\|^2 = 1$ often arises. Here z and $f(z)$ need not live in the same dimensional space. In one dimension, $|\zeta|$ denotes the magnitude of a complex number ζ.

Several copies of complex Euclidean spaces often arise in the same discussion. Suppose $N = K + L$. We will write $\mathbb{C}^N = \mathbb{C}^K \oplus \mathbb{C}^L$, where the symbol $\oplus$ denotes orthogonal sum. When A is a subspace of $\mathbb{C}^N$, we let $A^\perp$ denote the orthogonal complement of A, and thus $\mathbb{C}^N = A \oplus A^\perp$. In these settings, the Pythagorean theorem holds:

$$\|z \oplus w\|^2 = \|z\|^2 + \|w\|^2.$$

When $M < N$, we often regard $\mathbb{C}^M$ as a subspace of $\mathbb{C}^N$. For $w \in \mathbb{C}^M$, we write either $w \oplus 0$ or $(w, 0)$ for the corresponding element in $\mathbb{C}^N$.

We say a bit now about tensor products and discuss them in more detail in Chap. 2. If $v \in \mathbb{C}^M$ and $w \in \mathbb{C}^N$, we can form an element $v \otimes w$ in $\mathbb{C}^{MN}$ whose components are $v_j w_k$ for $1 \leq j \leq M$ and $1 \leq k \leq N$ in some specified order. It is easy to see that $\|v \otimes w\|^2 = \|v\|^2 \, \|w\|^2$; thus, the squared norm of a tensor product is the product of the squared norms. See Lemma 2.1.

Let Ω be an open subset of $\mathbb{C}^n$. Suppose $f : \Omega \to \mathbb{C}^N$ is a function. Then f is **holomorphic** if, for each $p \in \Omega$, f is complex differentiable at p. In other words, there is a (complex) linear map $df(p) : \mathbb{C}^n \to \mathbb{C}^N$ such that

$$f(p + h) = f(p) + df(p)h + \mathrm{error}(p, h),$$

where $\lim_{h \to 0} \frac{\mathrm{error}(p,h)}{\|h\|} = 0$. We assume the reader knows possible equivalent definitions; for example, f is locally given by a convergent (vector-valued) power series, or f satisfies the Cauchy-Riemann equations. On several occasions we will have holomorphic mappings defined on open balls, and we use without comment that the power series expansion converges uniformly on compact subsets of the ball.

Assume Ω is connected. If all the power series coefficients of f vanish at a point $p \in \Omega$, then f vanishes identically. The standard proof is to observe that the set of such points is both open and closed in Ω.

On occasion, we will need the following form of the maximum principle.

Proposition 1.1 (Maximum principle) *Let Ω be a bounded, open, and connected subset of $\mathbb{C}^n$. Suppose $f : \Omega \to \mathbb{C}^N$ is holomorphic and extends continuously to the boundary $\partial\Omega$. Then:*

- *For all $z \in \Omega$, $\|f(z)\| \leq \sup_{\partial\Omega} \|f\|$.*
- *If there exists a $z \in \Omega$ where $\|f(z)\| = \sup_{\partial\Omega} \|f\|$, then f is a constant.*

The following obvious consequence of the maximum principle provides a strong contrast to the real case. For example, the polynomial $(x_1)^2 + \cdots + (x_n)^2$ is constant on the unit sphere in $\mathbb{R}^n$ but it is not constant. A holomorphic function that is constant on too large of a set must itself be constant. We need only the following simple case.

Corollary 1.1 *Suppose f is holomorphic in a domain containing the closed unit ball, and f is constant on the sphere. Then f is constant.*

Proof If $z \mapsto f(z) - c$ vanishes on the sphere, then (by the first part of the maximum principle) $\|f(z) - c\|$ vanishes on the ball. Hence, $f - c$ vanishes identically on its domain. (Recall that a domain is connected.) $\square$

Recall that a linear transformation $L : \mathbb{C}^n \to \mathbb{C}^n$ is **unitary** if L preserves the inner product. Thus, $\langle Lz, Lw\rangle = \langle z, w\rangle$ for all z, w. It follows trivially that $\|Lz\|^2 = \|z\|^2$ for all z. The converse also holds. Thus, L is unitary if and only if $\|Lz\|^2 = \|z\|^2$ for all z. In other words, $\langle Lz, Lz\rangle = \langle z, z\rangle$ for all z implies the apparently stronger statement $\langle Lz, Lw\rangle = \langle z, w\rangle$ for all z, w. This fact provides an easy example of *polarization*.

Remark 1.2 (*Polarization*) It is often crucial in complex analysis to regard z and its conjugate $\overline{z}$ as independent variables. In rather general circumstances, when an identity involving z and $\overline{z}$ holds for all choices of z, we may replace $\overline{z}$ by w, and the identity will hold for all z, w. We will often have identities that hold on the unit sphere defined by $\|z\|^2 = 1$. We may then replace $\overline{z}$ in the identity with w and the resulting identity will hold when $\langle z, w\rangle = 1$.

Example 1.1 We give a beautiful example of polarization. Consider a harmonic function u defined near the origin in $\mathbb{R}^2$. We wish to find a holomorphic function f whose real part is u. Thus, we must have

$$\frac{f(z) + \overline{f(z)}}{2} = u\left(\frac{z + \overline{z}}{2}, \frac{z - \overline{z}}{2i}\right) = u(x, y). \tag{1.1}$$

We polarize (1.1) by assuming that z and $\overline{z}$ are independent variables. For example, if we suppose $f(0) = 0$ and set $\overline{z} = 0$, then we obtain the formula

$$f(z) = 2u\left(\frac{z}{2}, \frac{z}{2i}\right). \tag{1.2}$$

Formula (1.2) recovers f from u without the usual process involving differentiation, integration, and the Cauchy-Riemann equations. Careful thought shows that we are assuming here that u is real-analytic. Proving that a harmonic function in the plane is real-analytic is usually done by showing that it is the real part of a holomorphic function. The result about real-analyticity holds in all dimensions. It can be proved by estimating the size of the successive derivatives after starting with the mean-value property.

The following ideas are equivalent to polarization and arise throughout complex analysis. Consider a polynomial or convergent series

$$f(z, \overline{z}) = \sum_{a,b} c_{ab} z^a \overline{z}^b$$

in the variables $(z, \overline{z}) \in \mathbb{C}^n \times \mathbb{C}^n$ with $f(0, 0) = 0$. The terms in the series $\sum c_{a0} z^a$ are called **holomorphic terms**, those in $\sum c_{0b} \overline{z}^b$ are called **anti-holomorphic terms**, and the remaining terms are called **mixed terms**. One says **pure terms** to mean any terms that are not mixed. Suppose $f(z, \overline{z})$ vanishes identically. Then $c_{ab} = 0$ for all multi-indices a, b. In particular, the series of holomorphic terms, the series of anti-holomorphic terms, and the series of mixed terms all vanish identically.

Exercise 1.1 Use the technique of Example 1.1 to find a holomorphic function f whose real part is the given $u(x, y)$.

- $u(x, y) = x^2 - y^2$.
- $u(x, y) = e^x \cos(y)$.
- $u(x, y) = \ln(x^2 + y^2)$. (Note that we cannot set $\overline{z}$ equal to 0 here!)

Unitary transformations arise throughout this book. The definition implies that the composition of unitary maps is unitary, and the next proposition implies that the collection $\mathbf{U}(n)$ of unitary maps on $\mathbb{C}^n$ is a group. We can identify $\mathbf{U}(n)$, when regarded as unitary matrices, as a subset of complex Euclidean space of dimension n^2. An $n \times n$ matrix is unitary if and only if its column vectors form an orthonormal basis of $\mathbb{C}^n$. Thus, this subset is closed and bounded. The group operations of multiplication and taking inverses are smooth. The next two propositions summarize the basic facts about unitary transformations.

Proposition 1.2 *Let $L : \mathbb{C}^n \to \mathbb{C}^n$ be linear. The following are equivalent:*

(1) L is unitary.
(2) For all z, w, we have $\langle Lz, Lw \rangle = \langle z, w \rangle$.
(3) For all z, we have $\|Lz\|^2 = \|z\|^2$.
(4) L is invertible and $L^{-1} = L^*$. (Here, L^* is the adjoint of L.)

Proposition 1.3 *The unitary group $\mathbf{U}(n)$ is a compact Lie group.*

A **Lie group** G is a smooth manifold endowed with a binary operation $(g, h) \mapsto gh$ that makes G into a group, and for which this operation and the operation of taking inverses are smooth maps. We don't use any major results from the theory of Lie groups or their Lie algebras, but specific examples of Lie groups arise throughout the book. We mention some of the groups that will arise.

The special unitary group $\mathbf{SU}(n)$ consists of those unitary maps with determinant equal to 1. The n-torus $\mathbf{T}(n)$ consists of those diagonal maps

$$z = (z_1, ..., z_n) \mapsto U(z) = (e^{i\theta_1} z_1, ..., e^{i\theta_n} z_n).$$

We often identify the n-torus with $\mathbf{U}(1) \times \cdots \times \mathbf{U}(1)$. Notice also that operators permuting the variables are unitary. Hence, (by Cayley's theorem) every finite group is isomorphic to a subgroup of $\mathbf{U}(n)$.

Let $n = p + q$. The groups $\mathbf{U}(p, q)$ are groups of n-by-n matrices preserving the Hermitian form

$$\sum_{j=1}^{p} |z_j|^2 - \sum_{j=p+1}^{n} |z_j|^2.$$

The subgroup $\mathbf{SU}(p, q)$ consists of those matrices in $\mathbf{U}(p, q)$ with determinant 1. In this book, we use only $\mathbf{SU}(p, 1)$, for the following reason. The quotient of the group $\mathbf{SU}(n, 1)$ by its center is isomorphic to the group of holomorphic automorphisms of the unit ball in $\mathbb{C}^n$, but we prefer a more concrete description in terms of linear fractional transformations. See Sects. 2 and 3.

An interesting class of examples of finite unitary groups arises in this book. In each case, the group itself is cyclic of order p, but the representations as subgroups of $\mathbf{U}(2)$ differ.

Example 1.2 Let η be a primitive p-th root of unity and assume q is relatively prime to p. Let $\Gamma(p, q)$ denote the cyclic subgroup of $\mathbf{U}(2)$ generated by

$$\begin{pmatrix} \eta & 0 \\ 0 & \eta^q \end{pmatrix}.$$

The special cases $\Gamma(p, 1)$ and $\Gamma(2r + 1, 2)$ play surprising major roles in this book.

The next proposition has been used innumerable times by the author and will be applied on several occasions in this book.

Proposition 1.4 *Let B denote a ball in $\mathbb{C}^n$. Assume that $f : B \to \mathbb{C}^N$ and $g : B \to \mathbb{C}^M$ are holomorphic maps. Suppose $\|f(z)\|^2 = \|g(z)\|^2$ on B.*

- *If $M = N$, then there is a $U \in \mathbf{U}(N)$ such that $f(z) = Ug(z)$ for $z \in B$.*
- *If $M < N$, then there is a $U \in \mathbf{U}(N)$ such that $f(z) = U(g(z) \oplus 0)$.*

Proof See [15] or [19] for a proof. □

Exercise 1.2 Determine the real dimensions of $\mathbf{U}(n)$ and $\mathbf{SU}(n)$. (Proposition 1.7 and Corollary 1.4 give the answers when $n = 2$.)

We conclude this section by defining **rational sphere map**. The simplest examples are given by unitary maps. In Sect. 3, we find all the equi-dimensional examples (automorphisms of the unit ball when $n \geq 2$). Most of our discussion is devoted to rational sphere maps in the positive codimension case; in other words, when the target dimension exceeds the source dimension.

Definition 1.1 A *rational sphere map* is any rational map $f = \frac{p}{q}$ satisfying the following properties:

- $p : \mathbb{C}^n \to \mathbb{C}^N$ is a (holomorphic) polynomial.
- $q : \mathbb{C}^n \to \mathbb{C}$ is a (holomorphic) polynomial, with $q(z) \neq 0$ when $\|z\| \leq 1$.

- $f = \frac{p}{q}$ is reduced to lowest terms.
- $\|p(z)\|^2 = |q(z)|^2$ when $\|z\|^2 = 1$.

Let $p : \mathbb{C}^n \to \mathbb{C}^N$ be a polynomial. When p sends the unit sphere in the source to the unit sphere in the target, we call p a **polynomial sphere map**. When also each component of p is a single monomial, we call p a **monomial sphere map**.

Remark 1.3 By Proposition 1.4, we can draw a very strong conclusion if $\|p\|^2 = |q|^2$ holds on the *ball*. In fact, $p = L(q \oplus 0)$ for some linear map L. The equality $\|p\|^2 = |q|^2$ on the *sphere* is much weaker; we will spend most of this book studying its solutions!

2 The Groups $\mathbf{Aut}(\mathbb{B}_1)$, $\mathbf{SU}(2)$, and $\mathbf{SU}(1, 1)$

Before discussing the automorphism group of the unit ball in $\mathbb{C}^n$, we recall the situation in one dimension. We also briefly consider the special unitary group $\mathbf{SU}(2)$ and the group $\mathbf{SU}(1, 1)$.

Let $\text{Aut}(\mathbb{B}_1)$ denote the set of holomorphic maps $f : \mathbb{B}_1 \to \mathbb{B}_1$ such that f is bijective and has a holomorphic inverse. We begin with the famous Schwarz lemma.

Proposition 1.5 (Schwarz lemma) *Let $f : \mathbb{B}_1 \to \mathbb{C}$ be holomorphic. Suppose $f(0) = 0$ and $|f(z)| \le 1$. Then the stronger inequality $|f(z)| \le |z|$ holds on $\mathbb{B}_1$.*

Proof Since $f(0) = 0$, the function g defined by $g(z) = \frac{f(z)}{z}$ is also holomorphic. For any $r < 1$, its maximum absolute value on the disk $|z| \le r$ is achieved on the circle $|z| = r$. Therefore,

$$|g(z)| \le \max_{|z| \le r} \left(\frac{|f(z)|}{r} \right) \le \frac{1}{r}. \tag{1.3}$$

Letting r tend to 1 in (1.3) shows that $|g(z)| \le 1$ and hence $|f(z)| \le |z|$. □

Corollary 1.2 *Suppose $f : \mathbb{B}_1 \to \mathbb{B}_1$ is a holomorphic automorphism with $f(0) =$ 0. Then f is a rotation. Thus, $f(z) = e^{i\theta} z$.*

Proof Applying Proposition 1.5 to both f and f^{-1} gives $|f(z)| = |z|$. Either the one-dimensional version of Proposition 1.4 or elementary complex analysis forces $\frac{f(z)}{z}$ to be a constant map, and the conclusion follows. □

Lemma 1.1 *Suppose $|w| < 1$, $|\zeta| < 1$, and $e^{i\theta}$ is on the unit circle. Then*

- $|e^{i\theta} + \zeta \overline{w}|^2 = |1 + w e^{i\theta} \overline{\zeta}|^2$.
- $|e^{i\theta} w + \zeta|^2 < |e^{i\theta} + \zeta \overline{w}|^2$.

Proof By expanding the squared magnitudes, the equality is equivalent to

$$1 + 2\mathrm{Re}(e^{i\theta}\overline{\zeta}w) + |\zeta\overline{w}|^2 = 1 + 2\mathrm{Re}(e^{i\theta}(w\overline{\zeta})) + |w\overline{\zeta}|^2$$

and hence holds. Expanding the squared norms and cancelling the equal real part terms shows that the inequality is equivalent to $|w|^2 + |\zeta|^2 < 1 + |w\zeta|^2$, which is equivalent to $|w|^2(1 - |\zeta|^2) < 1 - |\zeta|^2$. This inequality holds because both $|w| < 1$ and $|\zeta| < 1$. □

Lemma 1.2 *For $|w| < 1$, $|\zeta| < 1$ put $f(z) = e^{i\theta}\frac{z-w}{1-\overline{w}z}$ and $g(z) = e^{i\phi}\frac{z-\zeta}{1-\overline{\zeta}z}$. Then their composition is of the same form:*

$$(g \circ f)(z) = e^{i\gamma}\frac{z - s}{1 - \overline{s}z}$$

where $e^{i\gamma} = e^{i\phi}\frac{e^{i\theta}+\zeta\overline{w}}{1+we^{i\theta}\overline{\zeta}}$ and $s = \frac{e^{i\theta}w+\zeta}{e^{i\theta}+\zeta\overline{w}}$. Furthermore, $|s| < 1$.

Proof Compute $g(f(z))$. Then clear denominators by multiplying by $1 - \overline{w}z$. Factor $(e^{i\theta} + \zeta\overline{w})$ from the numerator and $(1 + \overline{\zeta}we^{i\theta})$ from the denominator. The result is the claimed formula for s. Also, by Lemma 1.1, the factors we extracted have the same magnitude. Hence, $e^{i\phi}$ times their quotient is on the unit circle. Finally, note that $|s|^2 < 1$ if and only if $|e^{i\theta}w + \zeta|^2 < |e^{i\theta} + \zeta\overline{w}|^2$, which was proved in Lemma 1.1. □

Corollary 1.3 *The collection of maps in Lemma 1.2 form a group under composition.*

Proof By Lemma 1.2, the composition of such maps is of the same form. Putting $\zeta = -e^{i\theta}w$ shows that each such map has an inverse of the same form. □

Proposition 1.6 $\mathrm{Aut}(\mathbb{B}_1)$ *is the set of functions f for which*

$$f(z) = e^{i\theta}\frac{z - w}{1 - \overline{w}z}, \tag{1.4}$$

where $|w| < 1$ and $e^{i\theta}$ is on the unit circle.

Proof It follows from Corollary 1.3 that each such f is an automorphism. We must show that there are no others. Let h be an automorphism with $h(0) = w$. For f as in (1.4), $f \circ h$ is an automorphism sending w to 0. By Corollary 1.2 of Schwarz's lemma, $f \circ h$ is a rotation U, and $h = f^{-1} \circ U$ has the desired form. □

The following proposition will not be explicitly used in the book, but its importance in physics and geometry suggests its inclusion. See [38] and its references for the many uses of $\mathbf{SU}(2)$ in physics. It is useful for us because of the major role the unit sphere, especially in two complex dimensions, plays in this book. Furthermore, comparing $\mathbf{SU}(2)$ with $\mathbf{SU}(1,1)$ is interesting and $\mathbf{SU}(1,1)$ is closely related to $\mathrm{Aut}(\mathbb{B}_1)$.

Recall that sets are diffeomorphic if there is a smooth bijective map with a smooth inverse between them.

Proposition 1.7 *The sets* $\mathbf{SU}(2)$ *and the unit sphere* S^3 *are diffeomorphic.*

Proof Given $p = (z, w) \in \mathbb{C}^2$, we define $U(z, w)$ to be the 2-by-2 matrix

$$U(p) = U(z, w) = \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}.$$

The mapping U is obviously injective on all of $\mathbb{C}^2$. Assume $|z|^2 + |w|^2 = 1$. Computing $U(z, w)U(z, w)^*$ gives

$$\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \begin{pmatrix} \overline{z} & -w \\ \overline{w} & z \end{pmatrix} = \begin{pmatrix} |z|^2 + |w|^2 & 0 \\ 0 & |z|^2 + |w|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus, if $(z, w) \in S^3$, then $U(z, w)$ is unitary. Furthermore, $\det(U(z, w)) = 1$ as well. Therefore, $U : S^3 \to \mathbf{SU}(2)$. The map U is obviously smooth; we need to prove that it is bijective with a smooth inverse.

A 2-by-2 matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is unitary if these equations are met:

$$|a|^2 + |b|^2 = |a|^2 + |c|^2 = |c|^2 + |d|^2 = 1$$

$$a\overline{b} + c\overline{d} = 0.$$

It is in $\mathbf{SU}(2)$ if also $ad - bc = 1$. First suppose that $b = 0$. Then $|a| = 1$ and hence both $c = 0$ and $|d| = 1$. In this case, $M = U(a, 0)$. Suppose that $b \neq 0$; then $c = \frac{ad-1}{b}$ and we get the equation

$$a\overline{b} + \frac{ad - 1}{b}\overline{d} = 0.$$

It follows that $0 = a|b|^2 + (ad - 1)\overline{d} = a(1 - |d|^2) + a|d|^2 - \overline{d} = a - \overline{d}$ and hence $a = \overline{d}$. Thus, $c = -\overline{b}$ and $M = U(a, b)$. Hence, U is surjective. The inverse map T sending a unitary matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to its first row (a, b) is obviously also smooth. □

Corollary 1.4 *The sets* $S^3 \times S^1$ *and* $\mathbf{U}(2)$ *are diffeomorphic.*

Proof By the Proposition, it suffices to show that $\mathbf{SU}(2) \times S^1$ and $\mathbf{U}(2)$ are diffeomorphic. Given $M \in \mathbf{SU}(2)$ and $e^{i\theta} \in S^1$, put $A(M, e^{i\theta}) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & 1 \end{pmatrix} M$. Then A is a diffeomorphism. □

Proposition 1.7 implies that the unit sphere S^3 is a Lie Group. The only spheres that are Lie groups are S^1 and S^3, although this fact is not so easy to prove. Another useful result in physics (which is easy to prove) is that the quotient of $\mathbf{SU}(2)$ by the subgroup of two elements $\pm I$ is the special orthogonal group $\mathbf{SO}(3)$. For us, it will be important (and shown in Chapter 6) that all odd dimensional spheres have an unbounded realization associated with the Heisenberg group.

Exercise 1.3 What is the group multiplication on S^3 determined by Proposition 1.7? In other words, let $p_1 = (z_1, w_1)$ and $p_2 = (z_2, w_2)$ be points on the sphere. Define $p_1 * p_2$ by applying the inverse map T to the product $U(p_1)U(p_2)$. Thus, $p_1 * p_2 = T(U(p_1)U(p_2))$. Find this formula.

Exercise 1.4 In proving Corollary 1.4, what would go wrong if we defined A by $A(M, e^{i\theta}) = e^{i\theta} M$?

Let $n = p + q$. The groups $\mathbf{SU}(p, q)$ are groups of n-by-n matrices with determinant 1 and preserving the Hermitian form

$$\sum_{j=1}^{p} |z_j|^2 - \sum_{j=p+1}^{n} |z_j|^2.$$

In the special case where $n = 2$, we determine the relationship between $\mathbf{SU}(1, 1)$ and the automorphisms of the unit disk. Put $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. A matrix M lies in $\mathbf{SU}(1, 1)$ if and only if $\det(M) = 1$ and $M^* J M = J$. Assuming that $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we see, in a manner similar to the proof of Proposition 1.7, that $M \in \mathbf{SU}(1, 1)$ if and only if M has the form

$$M = \begin{pmatrix} a & b \\ \overline{b} & \overline{a} \end{pmatrix} \tag{1.5}$$

where now $|a|^2 - |b|^2 = 1$. Note the minus sign and that $a \neq 0$. Since $|a|^2 - |b|^2 = 1$ allows $|a|$ to be arbitrarily large, formula (1.5) shows that $\mathbf{SU}(1, 1)$ is not compact. Notice also that the *center* of $\mathbf{SU}(1, 1)$ consists of the matrices $\pm I$. (The center of a group G is the set of elements g such that $gh = hg$ for all $h \in G$.)

Let us identify a fraction $\frac{\zeta}{w}$ with the row vector $(\zeta \ w)$. Thus, the complex number z is identified with the row vector $(z \ 1)$. Given M in $\mathbf{SU}(1, 1)$, we multiply the row vector $(z \ 1)$ on the right by M to obtain the row vector $(az + \overline{b} \ \ bz + \overline{a})$. Our identification $\sim$ with this row vector as a fraction yields

$$(z\ \ 1)M \sim \frac{a}{\overline{a}}\left(\frac{z+\frac{\overline{b}}{a}}{1+\frac{b}{\overline{a}}z}\right). \tag{1.6}$$

Writing $\frac{a}{\overline{a}}$ in the form $e^{i\theta}$, we see that $(z\ \ 1)M \sim e^{i\theta}\psi(z)$ where ψ is a holomorphic automorphism of the unit disk. If we replace (a, b) by $(-a, -b)$, then the map on the right-hand side of (1.6) is unchanged. We obtain the following conclusion.

Proposition 1.8 *The group* $\text{Aut}(\mathbb{B}_1)$ *is isomorphic to the quotient of* $\mathbf{SU}(1, 1)$ *by its center* $\pm I$.

Proof Let M be as in (1.5). Define a map $T : \mathbf{SU}(1, 1) \to \text{Aut}(\mathbb{B}_1)$ by putting $T(M)$ to be the map on the right-hand side in (1.6). Note that $\frac{a}{\overline{a}}$ is an arbitrary element of the unit circle and $\frac{\overline{b}}{a}$ is an arbitrary element of the unit disk. Thus, by Proposition 1.6, T is surjective. Lemma 1.2 implies that T is a group homomorphism. It is not an isomorphism, because $T(-M) = T(M)$, but it is two-to-one. The conclusion follows. □

In dimension n, there is an isomorphism from $\mathbf{SU}(n, 1)$ divided by its center to $\text{Aut}(\mathbb{B}_n)$. We prefer the concrete expression, from Theorem 1.1 in the next section, in terms of the linear fractional transformations generalizing (1.4).

3 Automorphisms of the Unit Ball

Let Ω be a set and let $\text{Aut}(\Omega)$ denote the group of its automorphisms. Let us clarify what we mean by the term *automorphism*. An automorphism must be a bijection from Ω to itself. When Ω has some given algebraic structure, an automorphism must preserve this structure. An automorphism of a finite set is simply a permutation of that set. An automorphism of a vector space V is an invertible linear map from V to itself. For us, Ω will be an open subset of complex Euclidean space $\mathbb{C}^n$ and a **holomorphic automorphism** will be a biholomorphic mapping from Ω to itself. Thus, $f : \Omega \to \Omega$ is holomorphic (complex analytic), injective, and surjective. The inverse mapping f^{-1} is also holomorphic. Since the composition of functions is an associative operation, it follows that $\text{Aut}(\Omega)$ is a group under composition. When Ω is an open subset of some complex Euclidean space, the notation $\text{Aut}(\Omega)$ denotes the group of holomorphic automorphisms of Ω.

Remark 1.4 By definition, a mapping $f : \Omega \to \Omega$ is a holomorphic automorphism if $f : \Omega \to \Omega$ is holomorphic, injective, surjective, and the inverse mapping is also holomorphic. One can show that the holomorphicity of the inverse mapping follows automatically. By contrast, however, things differ for smooth functions. The function $x \mapsto x^3$ on $\mathbb{R}$ is of class C^∞, injective, and surjective, but the inverse function is not smooth at 0. For holomorphic maps, this situation does not arise.

Next, we give a concrete description of the automorphism group of the unit ball. The book [71] has a similar treatment and also contains a huge amount of analytic information about holomorphic functions on the unit ball.

Let $a \in \mathbb{C}^n$ satisfy $\|a\|^2 < 1$. Write $s = \sqrt{1 - \|a\|^2}$. Define a linear map L_a by

$$L_a(z) = \frac{\langle z, a \rangle}{s+1} a + sz. \tag{1.7}$$

We define a rational map ϕ_a by

$$\phi_a(z) = \frac{a - L_a z}{1 - \langle z, a \rangle}. \tag{1.8}$$

For each $a \in \mathbb{B}_n$, we will show that ϕ_a is an automorphism of the unit ball $\mathbb{B}_n$. Furthermore, ϕ_a is a rational sphere map.

Lemma 1.3 *For L_a as in (1.7), and ϕ_a as in (1.8), the following holds:*

- $L_a(a) = a$.
- $\langle L_a z, a \rangle = \langle z, a \rangle$.
- $L_a(L_a(z)) = \langle z, a \rangle a + (1 - \|a\|^2) z$.
- $\phi_a(a) = 0$.

Proof The first item:

$$L_a(a) = \frac{\langle a, a \rangle}{s+1} a + sa = \frac{\|a\|^2}{s+1} a + sa = \frac{1 - s^2}{1+s} a + sa = a.$$

The second item:

$$\langle L_a z, a \rangle = \langle \frac{\langle z, a \rangle a}{s+1}, a \rangle + s \langle z, a \rangle = \frac{\langle z, a \rangle}{s+1} (1 - s^2) + s \langle z, a \rangle = \langle z, a \rangle.$$

The third item:

$$L_a(L_a(z)) = L_a \left(\frac{\langle z, a \rangle}{s+1} a + sz \right) = \frac{\langle z, a \rangle}{s+1} L_a(a) + s \left(\frac{\langle z, a \rangle}{s+1} a + sz \right)$$

$$= \frac{\langle z, a \rangle}{s+1} (1+s) a + s^2 z = \langle z, a \rangle a + (1 - \|a\|^2) z.$$

The final item follows from the first item. □

Lemma 1.4 *ϕ_a maps the unit sphere to itself.*

Proof It suffices to show on the unit sphere that (1.9) holds:

$$\|a - L_a z\|^2 = \|a\|^2 - 2\mathrm{Re}\langle a, L_a z \rangle + \|L_a z\|^2 = |1 - \langle z, a \rangle|^2. \tag{1.9}$$

In anticipation of a later calculation we will show that

$$||a - L_a z||^2 - |1 - \langle z, a\rangle|^2 = (||a||^2 - 1)(1 - ||z||^2). \tag{1.10}$$

The conclusion follows from (1.10) upon setting $||z||^2 = 1$. To prove (1.10), we expand everything using $L_a(z) = \frac{\langle z,a\rangle}{s+1} + sz$ and the second item of Lemma 1.3. We also use $2\text{Re}\langle z, a\rangle = 2\text{Re}\langle a, z\rangle$. Putting these things together shows that the left-hand side of (1.10) equals

$$(||a||^2 - 1) + |\langle z, a\rangle|^2 \left(\frac{1 - s^2}{(s+1)^2} + \frac{2s}{s+1} - 1 \right) + s^2||z||^2 = (||a||^2 - 1) + s^2||z||^2.$$

Finally, we use $s^2 = 1 - ||a||^2$ to obtain (1.10). □

Lemma 1.5 $\phi_a \circ \phi_a = I$.

Proof After clearing denominators, we use the items in Lemma 1.3:

$$\phi_a(\phi_a(z)) = \frac{a - L_a\left(\frac{a - L_a(z)}{1 - \langle z,a\rangle}\right)}{1 - \langle \frac{a - L_a(z)}{1 - \langle z,a\rangle}, a\rangle}$$

$$= \frac{(1 - \langle z, a\rangle)a - L_a(a) + L_a(L_a(z))}{1 - \langle z, a\rangle - \langle a - L_a(z), a\rangle}$$

$$= \frac{(1 - ||a||^2)z}{(1 - ||a||^2)} = z.$$

□

Exercise 1.5 Write down an explicit formula for the automorphism ϕ_a when $a = (0, ..., 0, \alpha)$.

The automorphism group $\text{Aut}(\mathbb{B}_n)$ is the real Lie Group $\mathbf{SU}(n, 1)/Z$. Here Z denotes the *center* of the group $\mathbf{SU}(n, 1)$. See Exercise 1.6. Rather than regarding $\text{Aut}(\mathbb{B}_n)$ in this way, we give explicit formulas as rational mappings in Theorem 1.1. To better understand $\text{Aut}(\mathbb{B}_n)$, we need Corollary 1.5 below, an analogue of Schwarz's lemma (Proposition 1.5) in n-dimensions, and also Corollary 1.6.

Corollary 1.5 *Let $f : \mathbb{B}_n \to \mathbb{B}_N$ be holomorphic and $f(0) = 0$. Then $||f(z)|| \leq ||z||$ holds for $z \in \mathbb{B}_n$.*

Proof Choose a non-zero $z \in \mathbb{B}_n$. Let l be a linear functional on $\mathbb{C}^N$ with $l(f(z)) = ||f(z)|| \leq 1$ and $||l|| \leq 1$. Define the linear map $L : \mathbb{C} \to \mathbb{C}^N$ by $\zeta \mapsto \frac{\zeta z}{||z||}$. Then $||L|| = 1$ as well. Put $g = l \circ f \circ L$. Then g satisfies the hypotheses of Schwarz's lemma in one dimension and therefore $|g(\zeta)| \leq |\zeta|$. Put $\zeta = ||z||$. We get