Complex Function Theory

Second Edition

Donald Saranson

AMS AMERICAN MATHEMATICAL SOCIETY

Complex Function Theory

Second Edition

Donald Sarason

AMS
AMERICAN MATHEMATICAL SOCIETY

2000 *Mathematics Subject Classification.* Primary 30–01.

Front Cover: The figure on the front cover is courtesy of Andrew D. Hwang.

The Hindustan Book Agency has the rights to distribute this book in India, Bangladesh, Bhutan, Nepal, Pakistan, Sri Lanka, and the Maldives.

For additional information and updates on this book, visit
www.ams.org/bookpages/mbk-49

Library of Congress Cataloging-in-Publication Data

Sarason, Donald.
Complex function theory / Donald Sarason. — 2nd ed.
p. cm.
Includes index.
ISBN-13: 978-0-8218-4428-1 (alk. paper)
ISBN-10: 0-8218-4428-8 (alk. paper)
1. Functions of complex variables. I. Title.

QA331.7.S27 2007
515′.9—dc22 2007060552

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at `http://www.ams.org/`

10 9 8 7 6 5 4 3 2 1   12 11 10 09 08 07

Contents

Preface to the Second Edition

As is usual in a second edition, various minor flaws that had crept into the first edition have been corrected. Certain topics are now presented more clearly, it is hoped, after being rewritten and/or reorganized. The treatment has been expanded only slightly: there is now a section on division of power series, and a brief discussion of homotopy. (The latter topic was relegated to a couple of exercises in the first edition.) Four appendices have been added; they contain needed background which, experience has shown, is not possessed nowadays by all students taking introductory complex analysis.

In this edition, the numbers of certain exercises are preceded by an asterisk. The asterisk indicates that the exercise will be referred to later in the text. In many cases the result established in the exercise will be needed as part of a proof.

I am indebted to a number of students for detecting minor errors in the first edition, and to Robert Burckel and Bjorn Poonen for their valuable comments. Special thanks go to George Bergman and his eagle eye. George, while teaching from the first edition, read it carefully and provided a long list of suggested improvements, both in exposition and in typography. I owe my colleague Henry Helson, the publisher of the first edition, thanks for encouraging me to publish these *Notes* in the first place, and for his many kindnesses during our forty-plus years together at Berkeley.

The figures from the first edition have been redrawn by Andrew D. Hwang, whose generous help is greatly appreciated. I am indebted to Edward Dunne and his AMS colleagues for the patient and professional way they shepherded my manuscript into print.

Finally, as always, I am deeply grateful to my wife, Mary Jennings, for her constant support, in particular, for applying her TEXnical skills to this volume.

Berkeley, California
January 26, 2007

Preface to the First Edition[1]

These are the notes for a one-semester introductory course in the theory of functions of a complex variable. The aim of the notes is to help students of mathematics and related sciences acquire a basic understanding of the subject, as a preparation for pursuing it at a higher level or for employing it in other areas. The approach is standard and somewhat old-fashioned.

The user of the notes is assumed to have a thorough grounding in basic real analysis, as one can obtain, for example, from the book of W. Rudin cited in the list of references. Notions like metric, open set, closed set, interior, boundary, limit point, and uniform convergence are employed without explanation. Especially important are the notions of a connected set and of the connected components of a set. Basic notions from abstract algebra also occur now and then. These are all concepts that ordinarily are familiar to students by the time they reach complex function theory.

As these notes are a rather bare-bones introduction to a vast subject, the student or instructor who uses them may well wish to supplement them with other references. The notes owe a great deal to the book by L. V. Ahlfors and to the book by S. Saks and A. Zygmund, which, together with the teaching of George Piranian, were largely responsible for my own love affair with the subject. Several other excellent books are mentioned in the list of references.

[1] The first edition was published by Henry Helson under the title *Notes on Complex Function Theory*.

The notes contain only a handful of pictures, not enough to do justice to the strong geometric component of complex function theory. The user is advised to make his or her own sketches as an aid to visualization. Thanks go to Andrew Hwang for drawing the pictures.

The approach in these notes to Cauchy's theorem, the central theorem of the subject, follows the one used by Ahlfors, attributed by him to A. F. Beardon. An alternative approach based on Runge's approximation theorem, adapted from Saks and Zygmund, is also presented.

The terminology used in the notes is for the most part standard. Two exceptions need mention. Some authors use the term "region" specifically to refer to an open connected subset of the plane. Here the term is used, from time to time, in a less formal way. On the other hand, the term "contour" is used in the notes in a specific way not employed by other authors.

I wish to thank my Berkeley Math H185 class in the Spring Semester, 1994, for pointing out a number of corrections to the prepublication version of the notes, and my wife, Mary Jennings, who read the first draft of the notes and helped me to anticipate some of the questions students might raise as they work through this material. She also typed the manuscript. I deeply appreciate her assistance and support. The notes are dedicated to her.

Berkeley, California
June 8, 1994

Chapter I

Complex Numbers

The complex number system, $\mathbf{C}$, possesses both an algebraic and a topological structure. In algebraic terms $\mathbf{C}$ is a field: $\mathbf{C}$ is equipped with two binary operations, addition and multiplication, satisfying certain axioms (listed below). It is the field one obtains by adjoining to $\mathbf{R}$, the field of real numbers, a square root of -1. Remarkably, the field so created contains not only square roots of each of its elements, but even n-th roots for every positive integer n. All the more remarkable is that adjoining a solution of the single equation $x^2 + 1 = 0$ to $\mathbf{R}$ results in an algebraically closed field: every nonconstant polynomial with coefficients in $\mathbf{C}$ can be factored over $\mathbf{C}$ into linear factors. This is the "Fundamental Theorem of Algebra," first established by C. F. Gauss (1777–1855) in 1799. Many different proofs are now known—over one hundred by one estimate. Gauss himself discovered four. Despite the theorem's name, most of the proofs, including the simplest ones, are not purely algebraic. One of the standard ones is presented in Chapter VII.

When complex numbers were first introduced, in the 16th century, and for many years thereafter, they were viewed with suspicion, a feeling the reader perhaps has shared. In high school one learns how to add and multiply two complex numbers, $a+ib$ and $c+id$, by treating them as binomials, with the extra rule $i^2 = -1$ to be used in forming the product. According to this recipe,

$$(a+ib)+(c+id) = (a+c)+i\,(b+d)$$

$$(a+ib)(c+id) = (ac-bd)+i\,(ad+bc)\,.$$

Everything seems to work out, but where does this number i come from? How can one just "make up" a square root of -1?

Nowadays it is routine for mathematicians to create new number systems from existing ones using the basic constructions of set theory. We shall create **C** in a simple direct way by imposing an algebraic structure, suggested by the rules above for addition and multiplication, on $\mathbf{R}^2$, the set of ordered pairs of real numbers.

I.1. Definition of C

The complex number system is defined to be the set **C** of all ordered pairs (x, y) of real numbers equipped with operations of addition and multiplication defined as follows:

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2) \\ (x_1, y_1)(x_2, y_2) &= (x_1x_2 - y_1y_2,\ x_1y_2 + y_1x_2).\end{aligned}$$

I.2. Field Axioms

The operations of addition and multiplication on **C** *satisfy the following conditions (the field axioms).*

(i) $z_1 + z_2 = z_2 + z_1, \qquad z_1z_2 = z_2z_1$
(commutative laws for addition and multiplication)

(ii) $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3, \qquad z_1(z_2z_3) = (z_1z_2)z_3$
(associative laws for addition and multiplication)

(iii) $z_1(z_2 + z_3) = z_1z_2 + z_1z_3,$
(distributive law)

(iv) *The complex number* $(0, 0)$ *is an additive identity.*

(v) *Every complex number has an additive inverse.*

(vi) *The complex number* $(1, 0)$ *is a multiplicative identity.*

(vii) *Every complex number different from* $(0, 0)$ *has a multiplicative inverse.*

Properties (i), (iv), (v), (vi) and the first identity in (ii) follow easily from the definitions of addition and multiplication together with the properties of **R**. The verifications of the second identity in (ii) and of (iii) are straightforward but somewhat tedious; they are relegated to Exercise 1.2.1 below. As for (vii), if $(a, b) \neq (0, 0)$, then finding the multiplicative inverse (x, y) of (a, b) amounts to solving the pair of linear equations

$$ax - by = 1, \qquad bx + ay = 0$$

for x and y. The determinant of the system is $a^2 + b^2$, which is not 0, and thus a unique solution exists: $x = a/(a^2 + b^2)$, $y = -b/(a^2 + b^2)$. In other words,

$$(a,b)^{-1} = \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right).$$

Exercise I.2.1. Prove that $\mathbf{C}$ obeys the associative law for multiplication and the distributive law.

Exercise I.2.2. Find the multiplicative inverses of the complex numbers $(0,1)$ and $(1,1)$.

Exercise I.2.3. Think of $\mathbf{C}$ as a vector space over $\mathbf{R}$. Let $c = (a,b)$ be in $\mathbf{C}$, and regard multiplication by c as a real linear transformation T_c. Find the matrix M_c for T_c with respect to the basis $(1,0)$, $(0,1)$. Observe that the map $c \mapsto M_c$ preserves addition and multiplication. Conclude that the algebra of two-by-two matrices over $\mathbf{R}$ contains a replica of $\mathbf{C}$.

I.3. Embedding of R in C. The Imaginary Unit

The set of complex numbers of the form $(x,0)$ is a subfield of $\mathbf{C}$; it is the isomorphic image of $\mathbf{R}$ under the map $x \mapsto (x,0)$. Henceforth we shall identify this subfield with $\mathbf{R}$ itself; in particular, we shall notationally identify the complex number $(x,0)$ with the real number x. Additionally, we use the symbol i to denote the complex number $(0,1)$, the so-called imaginary unit; it is one of the two square roots of -1 $(= (-1,0))$, the other being $-i$. (The reader should verify these statements directly from the definition of how complex numbers multiply.) With these conventions, the complex number (x,y) can be written as $x+iy$ (or as $x+yi$).

I.4. Geometric Representation

Since a complex number is nothing but an ordered pair of real numbers, it can be envisioned geometrically as a point in the coordinatized Euclidean plane. To say it another way, each point in the plane can be labeled by a complex number. The real numbers correspond to points on the horizontal axis, which is thus referred to in this context as the real axis. Complex numbers of the form iy with y real, so-called purely imaginary numbers, correspond to points on the vertical axis, which is thus referred to as the imaginary axis. When wishing to emphasize this geometric interpretation, we shall refer to $\mathbf{C}$ as the "complex plane."

In geometric terms, the addition of two complex numbers is just vector addition according to the parallelogram law. The geometric interpretation of multiplication will be presented later, in Section I.9.

If $z = x+iy$ is a complex number, then the Euclidean distance of z from the origin is denoted by $|z|$ and is called the absolute value (or the modulus) of z: $|z| = \sqrt{x^2+y^2}$. The reflection of z with respect to the real axis is denoted by $\overline{z}$ and is called the complex conjugate of z: $\overline{z} = x - iy$. The coordinates of z on the horizontal and vertical axes are denoted by Re z and Im z, respectively, and are called the real and imaginary parts of z: $\operatorname{Re} z = x$, $\operatorname{Im} z = y$. The reader should verify the following basic identities:

$$|z| = |\overline{z}| = \sqrt{z\overline{z}}, \qquad \operatorname{Re} z = \tfrac{1}{2}(z+\overline{z}), \qquad \operatorname{Im} z = \tfrac{1}{2i}(z-\overline{z}).$$

Exercise* I.4.1. Prove that if z_1 and z_2 are complex numbers then $\overline{z_1+z_2} = \overline{z}_1 + \overline{z}_2$, $\overline{z_1 z_2} = \overline{z}_1\overline{z}_2$, and $|z_1 z_2| = |z_1||z_2|$.

Exercise* I.4.2. Prove that the nonzero complex numbers z_1 and z_2 are positive multiples of each other if and only if $z_1\overline{z}_2$ is real and positive. (Note that, in geometric terms, z_1 and z_2 are positive multiples of each other if and only if they lie on the same ray emanating from the origin.)

Exercise I.4.3. Prove that if a polynomial with real coefficients has the complex root z, then it also has $\overline{z}$ as a root.

Note: An exercise whose number is preceded by an asterisk will be referred to later in the text. In many cases the result established in the exercise will be needed as part of a proof.

I.5. Triangle Inequality

If z_1 and z_2 are complex numbers, then $|z_1+z_2| \le |z_1|+|z_2|$. The inequality is strict unless one of z_1 and z_2 is zero, or z_1 and z_2 are positive multiples of each other.

In view of the interpretation of addition in **C** as vector addition, the inequality expresses the geometric fact that the length of any side of a triangle does not exceed the sum of the lengths of the other two sides. To obtain an analytic proof we note that

$$\begin{aligned}(|z_1|+|z_2|)^2 - |z_1+z_2|^2 &= (|z_1|+|z_2|)^2 - (z_1+z_2)(\overline{z}_1+\overline{z}_2)\\ &= |z_1|^2+|z_2|^2+2|z_1||z_2| - z_1\overline{z}_1 - z_1\overline{z}_2 - z_2\overline{z}_1 - z_2\overline{z}_2\\ &= 2|z_1||\overline{z}_2| - z_1\overline{z}_2 - z_2\overline{z}_1\\ &= 2(|z_1\overline{z}_2| - \operatorname{Re} z_1\overline{z}_2).\end{aligned}$$

The reader will easily validate this string of equalities on the basis of some of the identities from the preceding section (including those in Exercise I.4.1). Now if z is any complex number it is plain from the basic definitions that $|z| \ge$ Re z, with equality if and only if z is real and nonnegative. From this we conclude, in view of the equality between the extreme left and right sides in the string above, that $(|z_1|+|z_2|)^2 \ge |z_1+z_2|^2$, with equality if and only if $z_1\overline{z}_2$ is real and nonnegative. By Exercise I.4.2 in the preceding section, the latter happens if and only if one of z_1 and z_2 is 0, or z_1 and z_2 are positive multiples of each other. This establishes the triangle inequality.

Exercise* I.5.1. Prove that if $z_1, z_2, \ldots, z_n$ are complex numbers then $|z_1 + z_2 + \cdots + z_n| \leq |z_1| + |z_2| + \cdots + |z_n|$.

Exercise* I.5.2. Prove that if z_1 and z_2 are complex numbers then $|z_1 - z_2| \geq |z_1| - |z_2|$. Determine the condition for equality.

I.6. Parallelogram Equality

If z_1 and z_2 are complex numbers, then

$$|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2\left(|z_1|^2 + |z_2|^2\right).$$

In geometric terms the equality says that the sum of the squares of the lengths of the diagonals of a parallelogram equals the sum of the squares of the lengths of the sides. To establish it we note that the left side equals

$$(z_1 + z_2)(\overline{z}_1 + \overline{z}_2) \; + \; (z_1 - z_2)(\overline{z}_1 - \overline{z}_2),$$

which, when multiplied out, becomes

$$z_1\overline{z}_1 + z_1\overline{z}_2 + z_2\overline{z}_1 + z_2\overline{z}_2 + z_1\overline{z}_1 - z_1\overline{z}_2 - z_2\overline{z}_1 + z_2\overline{z}_2.$$

After making the obvious cancellations, we obtain the right side.

I.7. Plane Geometry via Complex Numbers

The two preceding sections illustrate how geometric facts can be translated into the language of complex numbers, and vice versa. The following exercises contain further illustrations.

Exercise I.7.1. Prove that if the complex numbers z_1 and z_2 are thought of as vectors in $\mathbf{R}^2$ then their dot product equals Re $z_1\overline{z}_2$. Hence, those vectors are orthogonal if and only if $z_1\overline{z}_2$ is purely imaginary (equivalently, if and only if $z_1\overline{z}_2 + \overline{z}_1 z_2 = 0$).

Exercise I.7.2. Let z_1, z_2, z_3 be points in the complex plane, with $z_1 \neq z_2$. Prove that the distance from z_3 to the line determined by z_1 and z_2 equals

$$\frac{1}{2|z_2 - z_1|}\left|z_1(\overline{z}_2 - \overline{z}_3) + z_2(\overline{z}_3 - \overline{z}_1) + z_3(\overline{z}_1 - \overline{z}_2)\right|;$$

in particular, the points z_1, z_2, z_3 are collinear if and only if $z_1(\overline{z}_2 - \overline{z}_3) + z_2(\overline{z}_3 - \overline{z}_1) + z_3(\overline{z}_1 - \overline{z}_2) = 0$. (Suggestion: Reduce to the case where $z_1 = 0$ and z_2 is real and positive.)

Exercise I.7.3. Prove that if z_1, z_2, z_3 are noncollinear points in the complex plane then the medians of the triangle with vertices z_1, z_2, z_3 intersect at the point $\frac{1}{3}\left(z_1 + z_2 + z_3\right)$.

Exercise I.7.4. Prove that the distinct complex numbers z_1, z_2, z_3 are the vertices of an equilateral triangle if and only if

$$z_1^2 + z_2^2 + z_3^2 = z_1z_2 + z_2z_3 + z_3z_1.$$

I.8. C as a Metric Space

Since **C** as a set coincides with the Euclidean plane, $\mathbf{R}^2$, it becomes a metric space if we endow it with the Euclidean metric. In this metric, the distance between the two points z_1 and z_2 equals $|z_1 - z_2|$. We can thus use in **C** all of the standard notions from the theory of metric spaces, such as open and closed sets, compactness, connectedness, convergence, and continuity. For example, a sequence $(z_n)_1^\infty$ in **C** will be said to converge to the complex number z provided

$$\lim_{n\to\infty} |z - z_n| = 0.$$

Exercise I.8.1. Prove that each of the following inequalities defines an open subset of **C**:

- $|z| < 1$,
- $\mathrm{Re}z > 0$,
- $|z + z^2| < 1$.

Exercise I.8.2. Prove that the following complex-valued functions are continuous at those points of **C** where they are defined.

$$\text{(a) } f(z) = z^2, \qquad \text{(b) } g(z) = \frac{1}{z}, \qquad \text{(c) } h(z) = \frac{1}{z^2 - 1}.$$

Exercise* I.8.3. Prove that addition and multiplication define continuous maps of **C**×**C** into **C**.

I.9. Polar Form

Let $z = x + iy$ be a nonzero point in the complex plane, and let (r, θ) be polar coordinates for the point. Thus, $r = |z|$, and θ is the angle between the positive real axis and the ray from the origin through z, with the usual sign convention. The angle θ is called an argument of z, written $\theta = \arg z$. This angle is determined by z only to within addition of an integer multiple of 2π, in other words, if $\theta = \arg z$, then $\theta + 2\pi k = \arg z$ for every integer k. (This mild misuse of the equality sign will cause no trouble in practice.) The particular value of $\arg z$ in the interval $(-\pi, \pi]$ is called the principal value and denoted by $\mathrm{Arg}\, z$. For example, $\mathrm{Arg}\, 1 = 0$, $\mathrm{Arg}\, (-1) = \pi$, $\mathrm{Arg}\, i = \frac{\pi}{2}$, $\mathrm{Arg}\, (-i) = -\frac{\pi}{2}$.

Because $x = r\cos\theta$ and $y = r\sin\theta$, the number z can be rewritten as

$$z = r(\cos\theta + i\sin\theta).$$

The expression on the right is the polar form of z.

Consider two nonzero complex numbers in polar form,

$$z_1 = r_1(\cos\theta_1 + i\,\sin\theta_1) \text{ and } z_2 = r_2(\cos\theta_2 + i\,\sin\theta_2).$$

Their product is given by

$$z_1z_2 = r_1r_2[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i\,(\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2)],$$

which, because of the addition formulas for the sine and cosine functions, can be rewritten as

$$z_1z_2 = r_1r_2[\cos(\theta_1 + \theta_2) + i\,\sin(\theta_1 + \theta_2)].$$

The expression on the right side is in polar form, giving us a geometric picture of the effect of complex multiplication: when one multiplies two nonzero complex numbers, one multiplies their absolute values and adds their arguments.

Exercise* I.9.1. Prove that $\arg \overline{z} = \arg z^{-1} = -\arg z$ for any nonzero complex number z.

Exercise I.9.2. Let z_1, z_2, z_3 be distinct points on the unit circle (i.e., $|z_j| = 1$ for each j). Prove that

$$\arg\frac{z_1}{z_2} = 2\arg\frac{z_3 - z_1}{z_3 - z_2},$$

and interpret the equality geometrically. (Suggestion: $\arg a = 2\arg b$ if and only if $\overline{a}\,b^2$ is real and positive.)

I.10. De Moivre's Formula

If $z = r(\cos\theta + i\,\sin\theta)$ is a nonzero complex number then, for every integer n,

$$z^n = r^n\,(\cos n\theta + i\,\sin n\theta).$$

For $n > 0$, one can establish this by repeated applications of the product formula from the last section. The case $n = -1$ follows from Exercise I.9.1 in the last section. Once the case $n = -1$ has been established, the case $n < 0$ follows from the case $n > 0$.

A typical application: By de Moivre's formula,

$$\begin{aligned}\cos 3\theta + i\,\sin 3\theta &= (\cos\theta + i\,\sin\theta)^3\\ &= \cos^3\theta + 3\cos^2\theta\,(i\,\sin\theta) + 3\cos\theta\,(i\,\sin\theta)^2 + (i\,\sin\theta)^3\\ &= \cos^3\theta - 3\cos\theta\sin^2\theta + i\,(3\cos^2\theta\sin\theta - \sin^3\theta).\end{aligned}$$

Equating real and imaginary parts, we obtain the following trigonometric identities:

$$\cos 3\theta = \cos^3\theta - 3\cos\theta\sin^2\theta, \qquad \sin 3\theta = 3\cos^2\theta\sin\theta - \sin^3\theta.$$

Exercise I.10.1. Use de Moivre's formula to find expressions for $\cos 5\theta$ and $\sin 5\theta$ as polynomials in $\cos\theta$ and $\sin\theta$.

Exercise I.10.2. Establish the formulas

$$\cos\theta + i\,\sin\theta + 1 = 2\cos\frac{\theta}{2}\left(\cos\frac{\theta}{2} + i\,\sin\frac{\theta}{2}\right),$$
$$\cos\theta + i\,\sin\theta - 1 = 2i\,\sin\frac{\theta}{2}\left(\cos\frac{\theta}{2} + i\,\sin\frac{\theta}{2}\right).$$

Exercise I.10.3. Establish the formulas

$$\frac{1}{2} + \cos\theta + \cos 2\theta + \cdots + \cos n\theta = \frac{\sin(n+\frac{1}{2})\theta}{2\sin\frac{\theta}{2}},$$
$$\sin\theta + \sin 2\theta + \cdots + \sin n\theta = \frac{\cos\frac{\theta}{2} - \cos(n+\frac{1}{2})\theta}{2\sin\frac{\theta}{2}}$$

(important in the theory of Fourier series) by making the substitution $z = \cos\theta + i\,\sin\theta$ in the identity

$$1 + z + z^2 + \cdots + z^n = \frac{z^{n+1}-1}{z-1}, \qquad (z \neq 1).$$

I.11. Roots

De Moivre's formula enables us to find the n-th roots of any complex number. To illustrate, suppose we seek the cube roots of 1. Any such cube root obviously has absolute value 1, so it has the form $\cos\alpha + i\,\sin\alpha$ for some angle α. By de Moivre's formula we must have $\cos 3\alpha + i\,\sin 3\alpha = 1$, which is true if and only if $3\alpha = 2\pi k$ for some integer k. The cube roots of 1 are thus the numbers $\cos\frac{2\pi k}{3} + i\,\sin\frac{2\pi k}{3}$ with k an integer. Two different choices of k yield the same root if and only if the choices are congruent modulo 3. There are thus three distinct cube roots of 1, which we obtain, for example, from the choices $k = 0, 1, 2$; besides 1 itself, they are

$$\cos\frac{2\pi}{3} + i\,\sin\frac{2\pi}{3} = -\frac{1}{2} + \frac{i\sqrt{3}}{2},$$
$$\cos\frac{4\pi}{3} + i\,\sin\frac{4\pi}{3} = -\frac{1}{2} - \frac{i\sqrt{3}}{2}.$$

By similar reasoning one sees that if $z = r(\cos\theta + i\,\sin\theta)$ is any nonzero complex number and n is any positive integer, then the n-th roots of z are the numbers

$$r^{\frac{1}{n}}\left(\cos\frac{\theta + 2\pi k}{n} + i\,\sin\frac{\theta+2\pi k}{n}\right), \qquad k = 0, 1, \ldots, n-1.$$

Exercise I.11.1. Find all cube roots of i.

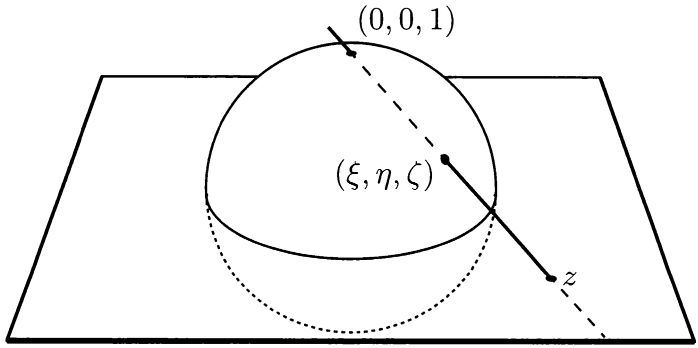

Figure 1. The stereographic projection

Exercise I.11.2. Find all eighth roots of 1.

Exercise I.11.3. An n-th root of 1 is called a primitive n-th root if it is not an m-th root of 1 for any positive integer m less than n. Prove that the number of primitive n-th roots of 1 is the same as the number of positive integers less than and relatively prime to n. Prove that one can obtain every n-th root of 1 by taking the first through the n-th powers of any given primitive n-th root.

Exercise I.11.4. Prove that the sum of the n-th roots of 1 equals 0, $(n > 1)$.

Exercise I.11.5. Let w be an n-th root of 1 different from 1 itself. Establish the formulas

$$1 + 2w + 3w^2 + \cdots + nw^{n-1} = \frac{n}{w-1},$$

$$1 + 4w + 9w^2 + \cdots + n^2 w^{n-1} = \frac{n^2}{w-1} - \frac{2n}{(w-1)^2}.$$

I.12. Stereographic Projection

To obtain another geometric picture of $\mathbf{C}$, we identify $\mathbf{C}$ with the horizontal coordinate plane in $\mathbf{R}^3$. Let $\mathbf{S}^2$ denote the unit sphere (the sphere of unit radius with center $(0,0,0)$) in $\mathbf{R}^3$. A given point $z = x + iy$ in $\mathbf{C}$ and the point $(0,0,1)$ (the north pole of $\mathbf{S}^2$) determine a line in $\mathbf{R}^3$. That line intersects $\mathbf{S}^2$ at one point other than the north pole. We denote that other point by (ξ, η, ζ) and map $\mathbf{C}$ to $\mathbf{S}^2$ by sending z to (ξ, η, ζ). The map is clearly one-to-one, and each point of the sphere other than the north pole is an image point.

To represent the map analytically we note that the line determined by z and $(0, 0, 1)$ is given parametrically by the function

$$(tx,\, ty,\, 1 - t), \qquad -\infty < t < \infty$$

($t = 0$ corresponds to the north pole and $t = 1$ to z). The point $(tx,\, ty,\, 1-t)$ lies on $\mathbf{S}^2$ provided

$$t^2x^2 + t^2y^2 + (1 - t)^2 = 1,$$

which has the two solutions $t = 0$ (giving the north pole) and $t = 2/(x^2 + y^2 + 1)$ (giving the image point (ξ, η, ζ)). From this one sees that our map is given by

$$z \mapsto \frac{(2\operatorname{Re} z,\, 2\operatorname{Im} z,\, |z|^2 - 1)}{|z|^2 + 1}.$$

The reader should verify that the inverse map is given by

$$(\xi, \eta, \zeta) \mapsto \frac{\xi + i\eta}{1 - \zeta}.$$

The latter map is the stereographic projection of cartographers.

Both the map from $\mathbf{C}$ to $\mathbf{S}^2$ and its inverse are continuous and so preserve topological properties such as convergence and compactness. For example, if a sequence in $\mathbf{C}$ converges, then so does the corresponding sequence on $\mathbf{S}^2$, and if a sequence on $\mathbf{S}^2$ converges to a point other than the north pole, then the corresponding sequence in $\mathbf{C}$ converges.

A few geometric features of the correspondence are easily seen from the figure above. For example, the unit circle in $\mathbf{C}$ corresponds to the equator of $\mathbf{S}^2$, the interior of the unit circle corresponds to the southern hemisphere of $\mathbf{S}^2$, and straight lines in $\mathbf{C}$ correspond to circles on $\mathbf{S}^2$ passing through the north pole.

In Exercise III.8.1 the reader will be asked to show that a circle on $\mathbf{S}^2$ not passing through the north pole corresponds under the stereographic projection to a circle in $\mathbf{C}$.

I.13. Spherical Metric

For z_1 and z_2 in $\mathbf{C}$, let $\rho(z_1, z_2)$ denote the Euclidean distance between the corresponding points on $\mathbf{S}^2$. This gives a metric in $\mathbf{C}$, the spherical metric, equivalent to the Euclidean metric (i.e., yielding the same family of open sets).

Exercise I.13.1. Establish the following formula for the spherical metric:

$$\rho(z_1, z_2) = \frac{2|z_1 - z_2|}{\sqrt{|z_1|^2 + 1}\,\sqrt{|z_2|^2 + 1}}.$$

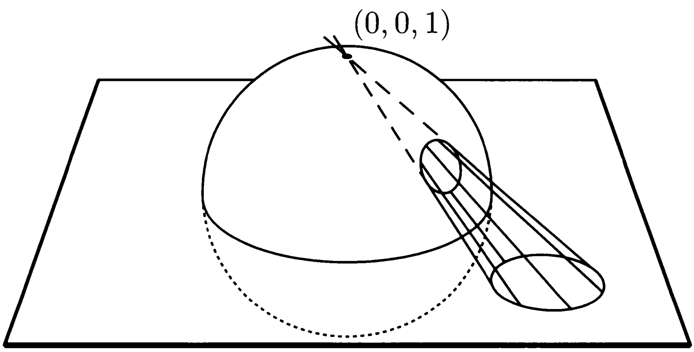

Figure 2. Transformation of circles under the stereographic projection

I.14. Extended Complex Plane

In many situations it is convenient to work with the enlargement of $\mathbf{C}$ that one obtains by appending a "point at infinity," corresponding to the north pole of $\mathbf{S}^2$. We denote this added point by ∞. (Set-theoretic purists can take ∞ to be any object not already in $\mathbf{C}$.) By the extended complex plane we shall mean the set $\overline{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$, to which we extend the spherical metric by defining $\rho(z, \infty)$, for z in $\mathbf{C}$, to be the Euclidean distance between the point on $\mathbf{S}^2$ corresponding to z and the north pole. Thus, for example, a sequence $(z_n)_1^\infty$ in $\mathbf{C}$ converges to ∞ in this metric if and only if $|z_n| \to \infty$ (in the standard sense).

The extended complex plane is often referred to as the Riemann sphere, after G. F. B. Riemann (1826–1866), a pioneer in our subject whose ideas had and continue to have a profound influence on its development.

Exercise I.14.1. Establish the formula

$$\rho(z, \infty) = \frac{2}{\sqrt{|z|^2 + 1}}.$$

Complex Differentiation

Having introduced the complex number system, we proceed to the development of the theory of functions of a complex variable, beginning with the notion of derivative. Although the definition of the derivative of a complex-valued function of a complex variable is formally the same as that of the derivative of a real-valued function of a real variable, the concept holds surprises, as we shall see.

Generally, we shall let z denote a variable point in the complex plane; its real and imaginary parts will be denoted by x and y, respectively.

II.1. Definition of the Derivative

Let the complex-valued function f be defined in an open subset G of $\mathbf{C}$. Then f is said to be differentiable (in the complex sense) at the point z_0 of G if the difference quotient $\dfrac{f(z)-f(z_0)}{z-z_0}$ has a finite limit as z approaches z_0. That limit is then called the derivative of f at z_0 and denoted by $f'(z_0)$:

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z)-f(z_0)}{z-z_0}.$$

What does the last equality mean? In ϵ–δ language, it is the statement that for every positive number ϵ there is a positive number δ such that

$$\left| \frac{f(z)-f(z_0)}{z-z_0} - f'(z_0) \right| < \epsilon$$

whenever $0 < |z - z_0| < \delta$.

As in calculus, the "d-notation" is also used for the derivative: if f is differentiable at z_0, then $f'(z_0)$ is according to convenience denoted alternatively by $\dfrac{df(z_0)}{dz}$.

NB. According to our definition, $f'(z_0)$ cannot be defined unless z_0 belongs to an open set in which f is defined.

II.2. Restatement in Terms of Linear Approximation

Let the complex-valued function f be defined in an open subset of $\mathbf{C}$ containing the point z_0. Then f is differentiable in the complex sense at z_0 if and only if there is a complex number c such that the function $R(z) = f(z) - f(z_0) - c(z - z_0)$ satisfies $\lim_{z\to z_0} \dfrac{R(z)}{z - z_0} = 0$, in which case $f'(z_0) = c$.

The statement is obvious in view of the equality

$$\frac{R(z)}{z - z_0} = \frac{f(z) - f(z_0)}{z - z_0} - c.$$

The statement says that f is differentiable at z_0, with $f'(z_0) = c$, if and only if f is well approximated near z_0 by the linear function $f(z_0) + c(z - z_0)$, in the sense that the remainder $R(z)$ in the approximation is small compared to the distance from z_0.

II.3. Immediate Consequences

The following properties of complex differentiation are proved from the basic definition in exactly the same way as the corresponding properties in the theory of functions of a real variable.

(i) *If f is differentiable at z_0 then f is continuous at z_0.*

(ii) *If f and g are differentiable at z_0, then $f + g$ and fg also are, and*

$$(f + g)'(z_0) = f'(z_0) + g'(z_0) \qquad \textit{(sum rule)};$$

$$(fg)'(z_0) = f'(z_0)g(z_0) + f(z_0)g'(z_0) \qquad \textit{(product rule)}.$$

If in addition $g(z_0) \neq 0$, then f/g is differentiable at z_0, and

$$\left(\frac{f}{g}\right)'(z_0) = \frac{f'(z_0)g(z_0) - f(z_0)g'(z_0)}{g(z_0)^2} \qquad \textit{(quotient rule)}.$$

(iii) *If f is differentiable at z_0 and g is differentiable at $f(z_0)$, then the composite function $g \circ f$ is differentiable at z_0 and*

$$(g \circ f)'(z_0) = g'(f(z_0))f'(z_0) \qquad \textit{(chain rule)}.$$

The proofs are left to the reader.

Exercise II.3.1. Prove statements (i)–(iii) in detail.

II.4. Polynomials and Rational Functions

From the definition of derivative it is immediate that a constant function is differentiable everywhere, with derivative 0, and that the identity function (the function $f(z) = z$) is differentiable everywhere, with derivative 1. Just as in elementary calculus one can show from the last statement, by repeated applications of the product rule, that, for any positive integer n, the function $f(z) = z^n$ is differentiable everywhere, with derivative nz^{n-1}. This, in conjunction with the sum and product rules, implies that every polynomial is everywhere differentiable: If $f(z) = c_n z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0$, where $c_0, \ldots, c_n$ are complex constants, then $f'(z) = nc_n z^{n-1} + (n-1)c_{n-1}z^{n-2} + \cdots + c_1$.

A function of the form f/g, where f and g are polynomials, is called a rational function. Such a function is defined wherever its denominator, g, does not vanish, hence everywhere except on a finite set. The quotient rule and the differentiability of polynomials imply that a rational function is differentiable at every point where it is defined and that its derivative is a rational function.

II.5. Comparison Between Differentiability in the Real and Complex Senses

Recall that a real-valued function u defined in an open subset G of $\mathbf{R}^2$ is said to be differentiable (in the real sense) at the point (x_0, y_0) of G if there are real numbers a and b such that the function $R(x,y) = u(x,y) - u(x_0,y_0) - a(x - x_0) - b(y - y_0)$ satisfies

$$\lim_{(x,y)\to(x_0,y_0)} \frac{R(x,y)}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} = 0.$$

In that case u has first partial derivatives at (x_0, y_0) given by

$$\frac{\partial u(x_0,y_0)}{\partial x} = a, \quad \frac{\partial u(x_0,y_0)}{\partial y} = b.$$

The reader will find this notion discussed in any multivariable calculus book.

To facilitate a comparison with complex differentiation, we restate the preceding definition in complex notation: the real-valued function u in the open subset G of $\mathbf{C}$ is by definition differentiable at the point $z_0 = x_0 + iy_0$ of G if there are real numbers a and b such that the function $R(z) = u(z) - u(z_0) - a(x - x_0) - b(y - y_0)$ satisfies

$$\lim_{z\to z_0} \frac{R(z)}{z - z_0} = 0.$$

Now suppose that f is a complex-valued function defined in the open subset G of $\mathbf{C}$, and let u and v denote its real and imaginary parts: $f = u+iv$. Given a point $z_0 = x_0 + iy_0$ of G and a complex number $c = a + ib$, we can write

$$\begin{aligned} R(z) = f(z) - f(z_0) - c(z - z_0) &= [u(z) - u(z_0) - a(x - x_0) + b(y - y_0)] \\ &\quad + i\,[v(z) - v(z_0) - b(x - x_0) - a(y - y_0)] \\ &= R_1(z) + iR_2(z). \end{aligned}$$

Clearly, $\lim\limits_{z\to z_0} \dfrac{R(z)}{z - z_0} = 0$ if and only if $\lim\limits_{z\to z_0} \dfrac{R_1(z)}{z - z_0} = 0$ and $\lim\limits_{z\to z_0} \dfrac{R_2(z)}{z - z_0} = 0$.

Referring to II.2, we can draw the following conclusion:

The function f is differentiable (in the complex sense) at z_0 if and only if u and v are differentiable (in the real sense) at z_0 and their first partial derivatives satisfy the relations $\dfrac{\partial u(z_0)}{\partial x} = \dfrac{\partial v(z_0)}{\partial y}$, $\dfrac{\partial u(z_0)}{\partial y} = -\dfrac{\partial v(z_0)}{\partial x}$. In that case,

$$f'(z_0) = \frac{\partial u(z_0)}{\partial x} + i\frac{\partial v(z_0)}{\partial x} = \frac{\partial v(z_0)}{\partial y} - i\frac{\partial u(z_0)}{\partial y}.$$

II.6. Cauchy-Riemann Equations

The two partial differential equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

are called the Cauchy-Riemann equations for the pair of functions u, v. As seen above, the equations are satisfied by the real and imaginary parts of a complex-valued function at each point where that function is differentiable.

Exercise II.6.1. At which points are the following functions f differentiable?

$$\text{(a) } f(z) = x, \qquad \text{(b) } f(z) = \overline{z}, \qquad \text{(c) } f(z) = \overline{z}^2.$$

Exercise II.6.2. Prove that the function $f(z) = \sqrt{|xy|}$ is not differentiable at the origin, even though it satisfies the Cauchy-Riemann equations there.

Exercise* II.6.3. Prove that the Cauchy-Riemann equations in polar coordinates are

$$r\frac{\partial u}{\partial r} = \frac{\partial v}{\partial \theta}, \qquad \frac{\partial u}{\partial \theta} = -r\frac{\partial v}{\partial r}.$$

II.7. Sufficient Condition for Differentiability

A theorem from the theory of functions of a real variable states that if a real-valued function of several variables has first partial derivatives, then it is differentiable at every point where those partial derivatives are continuous. (This can be found in any multivariable calculus book. For the convenience of readers who have not seen a proof, one is given in Appendix 1.) In combination with the necessary and sufficient condition from II.5, this gives the following useful sufficient condition for complex differentiability: *Let the complex-valued function $f = u + iv$ be defined in the open subset G of $\mathbf{C}$, and assume that u and v have first partial derivatives in G. Then f is differentiable at each point where those partial derivatives are continuous and satisfy the Cauchy-Riemann equations.*

II.8. Holomorphic Functions

A complex-valued function that is defined in an open subset G of $\mathbf{C}$ and differentiable at every point of G is said to be holomorphic (or analytic) in G. The simplest examples are polynomials, which are holomorphic in $\mathbf{C}$, and rational functions, which are holomorphic in the regions where they are defined. Later we shall see that the elementary functions of calculus—the exponential function, the logarithm function, trigonometric and inverse trigonometric functions, and power functions—all have complex versions that are holomorphic functions.

By II.5 we know that the real and imaginary parts of a holomorphic function have partial derivatives of first order obeying the Cauchy-Riemann equations. In the other direction, by II.7, if the real and imaginary parts of a complex-valued function have continuous first partial derivatives obeying the Cauchy-Riemann equations, then the function is holomorphic.

The asymmetry in the two preceding statements—the inclusion of a continuity condition in the second but not in the first—relates to an interesting and subtle theoretical point. The derivative of a holomorphic function, as will be shown later (in Section VII.8), is also holomorphic, so that in fact a holomorphic function is differentiable to all orders, and its real and imaginary parts have continuous partial derivatives to all orders. We shall only be able to prove this, however, after developing a fair amount of machinery. Meanwhile, we shall have to skirt around it occasionally.

Although, as we have seen above, some of the basic properties of real and complex differentiability are formally identical, the repeated differentiability of holomorphic functions points to a glaring dissimilarity. There are well-known examples of continuous real-valued functions on $\mathbf{R}$ that are nowhere differentiable. An indefinite integral of such a function is differentiable everywhere while its derivative is differentiable nowhere. By taking an n-fold indefinite integral, one can produce a function that is differentiable to order n yet whose n-th derivative is nowhere differentiable. Such "pathology" does not occur in the realm of complex differentiation.

From the basic rules of differentiation noted in Section II.3 one sees that if f and g are holomorphic functions defined in the same open set G, then $f+g$ and fg are also holomorphic in G, and f/g is holomorphic in $G\backslash g^{-1}(0)$. If f is holomorphic in G and g is holomorphic in an open set containing $f(G)$, then the composite function $g \circ f$ is holomorphic in G.

Exercise* II.8.1. Let the function f be holomorphic in the open disk D. Prove that each of the following conditions forces f to be constant: (a) $f' = 0$ throughout D; (b) f is real-valued in D; (c) $|f|$ is constant in D; (d) $\arg f$ is constant in D.

Exercise* II.8.2. Let the function f be holomorphic in the open set G. Prove that the function $g(z) = \overline{f(\overline{z})}$ is holomorphic in the set $G^* = \{\overline{z} : z \in G\}$.

II.9. Complex Partial Differential Operators

The partial differential operators $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ are applied to a complex-valued function $f = u + iv$ in the natural way:

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}, \quad \frac{\partial f}{\partial y} = \frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}.$$

We define the complex partial differential operators $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial \overline{z}}$ by

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \quad \frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

Thus, $\displaystyle\frac{\partial}{\partial x} = \frac{\partial}{\partial z} + \frac{\partial}{\partial \overline{z}}, \ \frac{\partial}{\partial y} = i\left(\frac{\partial}{\partial z} - \frac{\partial}{\partial \overline{z}}\right)$.

Intuitively one can think of a holomorphic function as a complex-valued function in an open subset of $\mathbf{C}$ that depends only on z, i.e., is independent of $\overline{z}$. We can make this notion precise as follows. Suppose the function $f = u + iv$ is defined and differentiable in an open set. One then has

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \frac{i}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right),$$

$$\frac{\partial f}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + \frac{i}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right).$$

The Cauchy-Riemann equations thus can be written $\frac{\partial f}{\partial \overline{z}} = 0$. As this is the condition for f to be holomorphic, it provides a precise meaning for the statement: "A holomorphic function is one that is independent of $\overline{z}$." If f is holomorphic, then (not surprisingly) $f' = \frac{\partial f}{\partial z}$, as the following calculation shows:

$$f' = \frac{\partial f}{\partial x} = \frac{\partial f}{\partial z} + \frac{\partial f}{\partial \overline{z}} = \frac{\partial f}{\partial z}.$$

II.10. Picturing a Holomorphic Function

One can visualize a real-valued function of a real variable by means of its graph, which is a curve in $\mathbf{R}^2$. A complex-valued function of a complex variable also has a graph, but its graph is a two-dimensional object in the four-dimensional space $\mathbf{C}\times\mathbf{C}$, something ordinary mortals cannot easily visualize. A more sensible approach, if one wants to obtain a geometric picture of a holomorphic function, is to think of the function as a map from the complex plane to itself, and to try to understand how the map deforms the plane; for example, how does it transform lines and circles?

The simplest case is that of a linear function, a function f of the form $f(z) = az + b$, where a and b are complex numbers, and $a \neq 0$ (to exclude the trivial case of a constant function). The map $z \mapsto az + b$ can be written as the composite of three easily understood transformations:

$$z \;\mapsto\; |a|z \;\mapsto\; az \;\mapsto\; az + b.$$

The first transformation in the chain is a scaling with respect to the origin by the factor $|a|$, a so-called homothetic map about the origin. The second transformation is multiplication by the number $a/|a|$, which is just rotation about the origin by the angle $\arg a$. The last transformation is translation by the vector b. We see in particular that the linear function $f(z) = az + b$ maps straight lines onto straight lines and preserves the angles between intersecting lines.

Linear functions are very special, but remember that a holomorphic function is a function that is well approximated locally by linear functions. If the function f is holomorphic in a neighborhood of the point z_0, one would expect it to behave near z_0 approximately like the linear function $z \mapsto f'(z_0)(z - z_0) + f(z_0)$. If $f'(z_0) = 0$ this will tell us little, but if $f'(z_0) \neq 0$ it should say something about the "infinitesimal" deformation produced by f near z_0. As we shall see, this is indeed the case: if $f'(z_0) \neq 0$, the holomorphic function f preserves the angles between curves intersecting at z_0. To make this precise we need some preliminaries about curves in the complex plane.

II.11. Curves in C

By a curve in **C** we shall mean a continuous function γ that maps an interval I of **R** into **C**. Thus, curves for us will always be parametrized curves. However, we shall often speak of curves as if they were subsets of **C**. For example, we shall say that the curve γ is contained in a given region of **C** if the range of γ is contained in that region.

Here are a few simple examples.

1. $\gamma(t) = (1-t)z_1 + tz_2 \ (-\infty < t < \infty)$.

Here, z_1 and z_2 are distinct points of **C**. This curve is a parametrization of the straight line determined by z_1 and z_2, the direction of the parametrization being from z_1 to z_2.

2. $\gamma(t) = \cos t + i\sin t \ (0 \le t \le 2\pi)$.

This is a parametrization of the unit circle, the circle being traversed once in the counterclockwise direction as t moves from the initial to the terminal point of the parameter interval $[0, 2\pi]$.

3. $\gamma(t) = \cos t - i\sin t \ (-2\pi \le t \le 2\pi)$.

This also is a parametrization of the unit circle, but this time the circle is traversed twice in the clockwise direction.

4. In this example, $\gamma(t)$ is defined piecewise:

$$\gamma(t) = \begin{cases} t, & 0 \le t \le 1, \\ 1 + (t-1)i, & 1 \le t \le 2, \\ i + 3 - t, & 2 \le t \le 3, \\ (4-t)i, & 3 \le t \le 4. \end{cases}$$

This is a parametrization of the square with vertices $0,\ 1,\ 1+i,\ i$. The square is traversed once in the counterclockwise direction.

The curve $\gamma : I \longrightarrow \mathbf{C}$ is said to be differentiable at the point t_0 of I if its real and imaginary parts are differentiable at t_0, or, what is equivalent, if the difference quotient $\dfrac{\gamma(t) - \gamma(t_0)}{t - t_0}$ approaches a finite limit as t tends to t_0. That limit is then denoted by $\gamma'(t_0)$. The curve γ is called differentiable if it is differentiable at each of its points; it is said to be of class C^1 if it is differentiable and its derivative, γ', is continuous.

The curve γ is said to be regular at the point t_0 if it is differentiable at t_0 and $\gamma'(t_0) \ne 0$. If γ is of class C^1 and regular at each point of its interval of definition, we call it a regular curve. The curves in the first three examples above are regular. The one in the fourth example is regular except at the points $1, 2, 3$ of the parameter interval $[0, 4]$.

A curve γ has a well-defined direction at each point t_0 where it is regular, namely, the direction determined by the derivative $\gamma'(t_0)$, referred to as the tangent direction. We can describe that direction, for example, by specifying the argument of $\gamma'(t_0)$, or by specifying the unit tangent vector, $\dfrac{\gamma'(t_0)}{|\gamma'(t_0)|}$.

Suppose γ_1 and γ_2 are two curves in $\mathbf{C}$, and suppose they have a point of intersection, say $\gamma_1(t_1) = \gamma_2(t_2)$. Suppose further that γ_j is regular at t_j, $j = 1,\ 2$. Then by the angle between γ_1 and γ_2 we shall mean the angle $\arg\gamma_2'(t_2) - \arg\gamma_1'(t_1)$ $(= \arg\gamma_2'(t_2)\overline{\gamma_1'(t_1)})$. In geometric terms, this is the angle through which one must rotate the unit tangent vector to γ_1 at t_1 to make it coincide with the unit tangent vector to γ_2 at t_2. Note that the angle depends on the order in which we take γ_1 and γ_2; reversal of the order leaves the magnitude of the angle the same but changes its sign. (To be completely precise, perhaps we should speak of the "angle between γ_1 and γ_2 corresponding to the parameter values t_1 and t_2" because the two curves might intersect for other pairs of parameter values. This degree of precision would not be worth the awkwardness of expression it would entail.)

Suppose that f is a holomorphic function in an open set G and that γ is a curve in G. Then we can apply f to γ to obtain the curve $f \circ \gamma$. Suppose γ is differentiable at t_0, and let $z_0 = \gamma(t_0)$. Then the standard argument justifying the chain rule applies to show that $f \circ \gamma$ is differentiable at t_0 and that $(f \circ \gamma)'(t_0) = f'(z_0)\gamma'(t_0)$. (Details are in Appendix 2.) Thus, if γ is regular at t_0 and if $f'(z_0) \neq 0$, then $f \circ \gamma$ is regular at t_0, and one obtains the direction of $f \circ \gamma$ at t_0 from that of γ at t_0 by adding $\arg f'(z_0)$.

II.12. Conformality

Let f be a holomorphic function defined in the open subset G of $\mathbf{C}$, and let z_0 be a point of G such that $f'(z_0) \neq 0$. Let γ_1 and γ_2 be curves such that $\gamma_1(t_1) = \gamma_2(t_2) = z_0$, and such that γ_j is regular at t_j, $j = 1,\ 2$. Then the angle between $f \circ \gamma_1$ and $f \circ \gamma_2$ equals the angle between γ_1 and γ_2.

This statement follows immediately from the discussion preceding it, from which one sees that

$$\arg\,(f \circ \gamma_j)'(t_j) = \arg f'(z_0) + \arg\gamma_j'(t_j), \qquad j = 1,\ 2.$$

The function $f(z) = z^2$ shows what can happen if the hypothesis $f'(z_0) \neq 0$ is dropped. This function, whose derivative vanishes at the origin, transforms two lines through the origin making an angle α into two lines making an angle 2α. On the other hand, as we shall see in a later chapter, the derivative of a nonconstant holomorphic function can vanish only on an isolated set of points, so the angle-preservation property given by the theorem above is the rule rather than the exception.

A map from the plane to the plane is called conformal at the point z_0 if it preserves the angles between pairs of regular curves intersecting at z_0. Thus, we can restate II.11 by saying that a holomorphic function is conformal at each point where its derivative does not vanish.

II.13. Conformal Implies Holomorphic

We shall now show that conformal maps are necessarily holomorphic. We begin with the simplest case, that of a linear transformation of the plane. Linear here means linear as a transformation of the real vector space $\mathbf{R}^2$ to itself. If f is such a map then $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are constants, and f is uniquely determined by those constants plus the condition $f(0) = 0$. Letting $a = \frac{\partial f}{\partial z}$ and $b = \frac{\partial f}{\partial \bar{z}}$ (also constants), we see that $f(z) = az + b\bar{z}$. Suppose this map preserves the angles between pairs of directed lines intersecting at the origin. We shall then prove that $b = 0$. We may assume that $a + b \neq 0$ (since otherwise the transformation would send the whole real axis to the origin) and, that done, that $a \neq 0$ (since otherwise the map would be anticonformal—it would reverse the angles between pairs of directed lines). Let λ be a complex number of absolute value 1. Our map sends the real line to the directed line through the origin determined by $a+b$, and it sends the directed line through the origin determined by λ to the one determined by $a\lambda + b\bar{\lambda}$. Our assumption about angle preservation thus implies that

$$\arg(a\lambda + b\bar{\lambda}) \; - \; \arg(a+b) \; = \; \arg\lambda.$$

Since the left side in this equality equals

$$\arg\lambda + \arg\left(a + \frac{b\bar{\lambda}}{\lambda}\right) - \arg(a+b),$$

the equality reduces to

$$\arg\left(a + \frac{b\bar{\lambda}}{\lambda}\right) \; = \; \arg(a+b)\,.$$

Now, if $b \neq 0$ then, as λ traverses the unit circle, the point $a + \dfrac{b\bar{\lambda}}{\lambda}$ (twice) traverses the circle with center a and radius $|b|$, in violation of the preceding equality, which says that $a + \frac{b\overline{\lambda}}{\lambda}$ lies on the ray through the origin determined by $a+b$. We can conclude that $b = 0$; in other words, our map is given by $z \mapsto az$, a holomorphic function, as desired.

The preceding result will serve as a lemma for its generalization to C^1 maps. We consider a complex-valued function $f = u + iv$ defined on an open subset G of $\mathbf{C}$. We assume that u and v have continuous first partial derivatives. Then, if γ is a differentiable curve in G, the curve $f \circ \gamma$ is also differentiable. In fact, suppose z_0 is a point on γ, say $\gamma(t_0) = z_0$. Let

$$a = \frac{\partial f(z_0)}{\partial z}, \quad b = \frac{\partial f(z_0)}{\partial \overline{z}}.$$

An application of the chain rule, the details of which are in Appendix 2, then shows that

$$(f \circ \gamma)'(t_0) = a\gamma'(t_0) + b\overline{\gamma'(t_0)}.$$

Hence, if f preserves the angles between pairs of regular curves intersecting at z_0, then the linear map $z \mapsto az + b\overline{z}$ preserves the angles between pairs of directed lines through the origin. By what is established above, that means $b = 0$ and $a \neq 0$. The equality $b = 0$ just says that the functions u and v satisfy the Cauchy-Riemann equations at z_0, which, as noted in Section II.7, implies that f is differentiable (in the complex sense) at z_0. Moreover, $f'(z_0) = \frac{\partial f(z_0)}{\partial z} = a$.

The following theorem has been proved: *Let f be a complex-valued function, defined in an open subset G of $\mathbf{C}$, whose real and imaginary parts have continuous first partial derivatives. If f preserves the angles between regular curves intersecting in G, then f is holomorphic, and f' is never* 0.

II.14. Harmonic Functions

The complex-valued function f, defined in an open subset of $\mathbf{C}$, is said to be harmonic if it is of class C^2 and satisfies Laplace's equation:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

This equation and its higher-dimensional versions play central roles in many branches of mathematics and physics. Of course, the complex-valued function f is harmonic if and only if its real and imaginary parts are.

II.15. Holomorphic Implies Harmonic

A holomorphic function is harmonic, provided it is of class C^2.

As noted in Section II.8, we shall prove later that a holomorphic function is of class C^k for all k, at which point we can drop the proviso in the preceding statement. To establish the proposition, let the function $f = u + iv$ be holomorphic and of class C^2. By the Cauchy-Riemann equations, we have

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial v}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial v}{\partial x}\right) = -\frac{\partial^2 u}{\partial y^2},$$

which proves that u is harmonic. Similar reasoning proves the same result for v, and thus f is harmonic.

II.16. Harmonic Conjugates

The reasoning in the preceding section shows that a pair of real-valued C^2 functions u and v, defined in the same open subset of $\mathbf{C}$, will be harmonic if they satisfy the Cauchy-Riemann equations:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

In this situation one says that v is a harmonic conjugate of u. Phrased differently, if u and v are real valued and of class C^2, then v is a harmonic conjugate of u if and only if $u + iv$ is holomorphic.

Note that harmonic conjugates are not unique: if v is a harmonic conjugate of u then so are the functions that differ from v by constants. That is essentially the extent of nonuniqueness, as we shall see later (and, in a special case, in Exercise II.16.4 below). A natural question is whether every harmonic function has a harmonic conjugate. We shall eventually develop enough machinery to deal with this question.

Friendly Advice. When beginning the study of complex analysis and faced with a problem in the subject, the initial response of many students is to reduce the problem to one in real variables, a subject they have previously studied. Such a reduction can sometimes be helpful, but at other times it can make things overly complicated. (Some of the exercises below illustrate this point.) Try to get into the habit of "thinking complex."

Exercise II.16.1. For which values of the real constants a, b, c, d is the function $u(x,y) = ax^3 + bx^2y + cxy^2 + dy^3$ harmonic? Determine a harmonic conjugate of u in the cases where it is harmonic.

Exercise* II.16.2. Prove that Laplace's equation can be written in polar coordinates as

$$r^2\frac{\partial^2 u}{\partial r^2} + r\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial \theta^2} = 0.$$

Exercise II.16.3. Find all real-valued functions h, defined and of class C^2 on the positive real line, such that the function $u(x,y) = h(x^2+y^2)$ is harmonic.

Exercise II.16.4. Prove that, if u is a real-valued harmonic function in an open disk D, then any two harmonic conjugates of u in D differ by a constant.

Exercise II.16.5. Suppose that u is a real-valued harmonic function in an open disk D, and suppose that u^2 is also harmonic. Prove that u is constant.

Exercise II.16.6. Prove that if the harmonic function v is a harmonic conjugate of the harmonic function u, then the functions uv and $u^2 - v^2$ are both harmonic.

Exercise II.16.7. Prove (assuming equality of second-order mixed partial derivatives) that

$$\frac{\partial^2}{\partial \bar{z} \partial z} = \frac{1}{4}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right).$$

Thus, Laplace's equation can be written as $\dfrac{\partial^2 f}{\partial \bar{z} \partial z} = 0$.

Exercise II.16.8. Prove that if u is a real-valued harmonic function then the function $\dfrac{\partial u}{\partial z}$ is holomorphic.

Chapter III

Linear-Fractional Transformations

A linear-fractional transformation is a nonconstant function ϕ of the form $\phi(z) = \dfrac{az+b}{cz+d}$, where a, b, c, d are complex constants. These are the simplest nonconstant rational functions. Linear-fractional transformations have many interesting algebraic and geometric properties and they surface in many different aspects of complex function theory.

It is instructive to introduce linear-fractional transformations from a "projective" viewpoint.

III.1. Complex projective space

By $\mathbf{C}^2$ we shall mean the complex vector space of two-dimensional complex column vectors:

$$\mathbf{C}^2 = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} : z_1, z_2 \in \mathbf{C} \right\}.$$

We introduce an equivalence relation between nonzero vectors in $\mathbf{C}^2$ by declaring that two vectors shall be equivalent if they are linearly dependent, in other words, if they are (complex) scalar multiples of each other. The equivalence classes are thus the one-dimensional subspaces of $\mathbf{C}^2$, with the origin deleted. The equivalence class determined by the pair

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

will be denoted by $[z_1, z_2]$. The space of equivalence classes is called one-dimensional complex projective space and denoted by $\mathbf{CP}^1$.

There is a natural one-to-one map of $\mathbf{CP}^1$ onto $\overline{\mathbf{C}}$, namely, the map

$$[z_1, z_2] \mapsto \frac{z_1}{z_2},$$

the right side being interpreted as ∞ in case $z_2 = 0$. With this we have a new way of thinking about the extended complex plane.

III.2. Linear-fractional transformations

A nonsingular linear transformation of $\mathbf{C}^2$ to itself maps each one-dimensional subspace of $\mathbf{C}^2$ onto a one-dimensional subspace. It thus induces a map from $\mathbf{CP}^1$ to itself; that map is easily seen to be a bijection (one-to-one and onto). In view of the correspondence between $\mathbf{CP}^1$ and $\overline{\mathbf{C}}$, therefore, each nonsingular linear transformation of $\mathbf{C}^2$ determines a bijection of $\overline{\mathbf{C}}$ with itself.

The linear transformations of $\mathbf{C}^2$ are in one-to-one correspondence with the two-by-two complex matrices, each matrix corresponding to the transformation it induces via left multiplication. Suppose

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is such a matrix, and suppose that M is nonsingular, in other words, that $\det M = ad - bc \neq 0$. Let ϕ denote the bijection of $\overline{\mathbf{C}}$ that M induces according to the recipe above. If z is a point of $\mathbf{C}$, then

$$M \begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} az + b \\ cz + d \end{pmatrix},$$

from which we conclude that $\phi(z) = \dfrac{az+b}{cz+d}$. Thus, ϕ is a linear-fractional transformation. (We shall always regard these transformations as acting on $\overline{\mathbf{C}}$. If $c = 0$ one has $\phi(\infty) = \infty$, and otherwise one has $\phi(\infty) = \dfrac{a}{c}$ and $\phi\left(\dfrac{-d}{c}\right) = \infty$.)

Certain conclusions are immediate. First, the composite of two linear-fractional transformations is again a linear-fractional transformation. More precisely, if ϕ_1 and ϕ_2 are the linear-fractional transformations induced by the matrices M_1 and M_2, respectively, then $\phi_1 \circ \phi_2$ is the linear-fractional transformation induced by $M_1 M_2$. Since the identity transformation of $\overline{\mathbf{C}}$ is the linear-fractional transformation induced by the identity matrix, it follows that the inverse of a linear-fractional transformation is a linear-fractional transformation: if ϕ is the linear-fractional transformation induced by the matrix M, then ϕ^{-1} is induced by M^{-1}.

Thus, the linear-fractional transformations form a group under composition. (If the concept of a group is new to you, see Appendix 3.) Under the map that sends each matrix to its induced linear fractional transformation, this group is the homomorphic image of the group $GL(2,\mathbf{C})$, the general linear group for dimension two (alias the group of nonsingular two-by-two complex matrices). The kernel of the homomorphism consists of the nonzero scalar multiples of I_2, the two-by-two identity matrix. The group of linear-fractional transformations is thus isomorphic to the quotient group $GL(2,\mathbf{C})/(\mathbf{C}\backslash\{0\})I_2$.

III.3. Conformality

If ϕ is a linear-fractional transformation and ψ is its inverse, then the chain rule gives $\psi'(\phi(z))\phi'(z) = 1$ (provided $\phi(z) \neq \infty$), implying that ϕ' is never 0. The same conclusion is easily reached directly, for if $\phi(z) = \dfrac{az+b}{cz+d}$, then a short calculation gives

$$\phi'(z) = \frac{ad-bc}{(cz+d)^2}\,.$$

Hence, a linear-fractional transformation is a conformal map.

III.4. Fixed points

A linear-fractional transformation, except for the identity transformation, has one or two fixed points.

For the proof, suppose the transformation is $\phi(z) = \dfrac{az+b}{cz+d}$. It is evident that ∞ is a fixed point if and only if $c = 0$. The finite fixed points are the solutions for z of the equation $cz^2 + (d-a)z - b = 0$. If $c \neq 0$ that equation has either two distinct roots or one repeated root (given by the standard quadratic formula). If $c = 0$, then there is one root in case $d \neq a$ and none in case $d = a$ (since in the last case $b \neq 0$). In all cases, therefore, the number of fixed points is one or two.

III.5. Three-fold transitivity

If z_1, z_2, z_3 are three distinct points of $\overline{\mathbf{C}}$, and w_1, w_2, w_3 are three distinct points of $\overline{\mathbf{C}}$, then there is a unique linear-fractional transformation ϕ such that $\phi(z_j) = w_j$, $j = 1,\ 2,\ 3$.

To establish the uniqueness we note that if the two linear-fractional transformations ϕ and ψ both have the required property, then $\psi^{-1}\circ\phi$ is a linear-fractional transformation with three fixed points; hence it is the identity, by III.4, implying that $\phi = \psi$. This settles uniqueness.

As for existence, because the linear-fractional transformations form a group, it will suffice to prove that ϕ exists for some particular choice of w_1,

w_2, w_3. (See Appendix 3 for an explanation of this assertion.) We consider the case $w_1 = \infty$, $w_2 = 0$, $w_3 = 1$. It is then a simple matter to write down the required ϕ:

$$\phi(z) = \begin{cases} \frac{(z-z_2)(z_1-z_3)}{(z-z_1)(z_2-z_3)} & \text{if } z_1,\ z_2,\ z_3 \text{ are all finite,} \\ \frac{z-z_2}{z_3-z_2} & \text{if } z_1 = \infty, \\ \frac{z_3-z_1}{z-z_1} & \text{if } z_2 = \infty, \\ \frac{z-z_2}{z-z_1} & \text{if } z_3 = \infty\,. \end{cases}$$

This settles existence.

Exercise III.5.1. Exhibit the linear-fractional transformation that maps 0, 1, ∞ to 1, ∞, $-i$, respectively.

Exercise III.5.2. Given four distinct points z_1, z_2, z_3, z_4 in $\overline{\mathbf{C}}$, their cross ratio, which is denoted by $(z_1, z_2; z_3, z_4)$, is defined to be the image of z_4 under the linear-fractional transformation that sends z_1, z_2, z_3 to ∞, 0, 1, respectively. Prove that if ϕ is a linear-fractional transformation then $(\phi(z_1), \phi(z_2); \phi(z_3), \phi(z_4)) = (z_1, z_2; z_3, z_4)$.

III.6. Factorization

It was observed in Section II.9 that every linear function (i.e., linear-fractional transformation with constant denominator) can be written as the composite of a homothetic map, a rotation, and a translation. Those three kinds of transformations, plus the transformation $z \mapsto \frac{1}{z}$, called the inversion, are enough to build the general linear-fractional transformation. Indeed, if $\phi(z) = \frac{az+b}{cz+d}$, and $c \neq 0$, then ϕ can be rewritten as

$$\phi(z) = \frac{b - ad/c}{cz+d} + \frac{a}{c},$$

from which one sees that one can obtain ϕ by applying a linear function followed by the inversion followed by a translation. Because of the natural isomorphism between $GL(2, \mathbf{C})/(\mathbf{C}\backslash\{0\})I_2$ and the group of linear fractional transformations, it follows that every matrix in $GL(2, \mathbf{C})$ is a scalar multiple of a product of matrices of the following four kinds:

$$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}, \quad k > 0$$

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, \quad |\lambda| = 1$$

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad b \in \mathbf{C}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Exercise III.6.1. Prove, either directly or by using the last result, that every matrix in $GL(2, \mathbf{C})$ is a product of matrices of the preceding four kinds.

Exercise III.6.2. Two linear-fractional transformations ϕ_1 and ϕ_2 are said to be conjugate if there is a linear-fractional transformation ψ such that $\phi_2 = \psi \circ \phi_1 \circ \psi^{-1}$. Prove all translations, except for translation by 0, are mutually conjugate.

Exercise III.6.3. Prove a linear-fractional transformation with only one fixed point is conjugate to a translation.

Exercise III.6.4. Prove the linear-fractional transformations $\phi_1(z) = \frac{1}{z}$ and $\phi_2(z) = -z$ are conjugate.

III.7. Clircles

By a clircle we shall mean a circle in $\mathbf{C}$, or a straight line in $\mathbf{C}$ together with the point at ∞. According to Exercise III.8.1 below, clircles correspond under the stereographic projection to circles on $\mathbf{S}^2$.

III.8. Preservation of clircles

A linear-fractional transformation maps clircles onto clircles.

This is obviously true of homothetic maps, rotations, and translations. It will therefore suffice, by III.5, to prove it is true of the inversion. The equation of the general clircle, or rather, the part of it in $\mathbf{C}$, has the form

$$\alpha|z|^2 + \beta \operatorname{Re} z + \gamma \operatorname{Im} z + \delta = 0,$$

where α, β, γ, δ are real constants. Under the map $z \mapsto \dfrac{1}{z}$, the locus of the preceding equation is sent to the locus of the equation

$$\frac{\alpha}{|z|^2} + \frac{\beta \operatorname{Re} z}{|z|^2} - \frac{\gamma \operatorname{Im} z}{|z|^2} + \delta = 0,$$

which can be rewritten as

$$\delta |z|^2 + \beta \operatorname{Re} z - \gamma \operatorname{Im} z + \alpha = 0,$$

again the equation of a clircle. This establishes the proposition.

Exercise III.8.1. Prove that the stereographic projection maps a circle on $\mathbf{S}^2$ onto a clircle, and vice versa.

Exercise III.8.2. Prove that the four distinct points z_1, z_2, z_3, z_4 of $\overline{\mathbf{C}}$ lie on a clircle if and only if the cross ratio $(z_1, z_2; z_3, z_4)$ is real.

III.9. Analyzing a linear-fractional transformation—an example

A clircle in $\overline{\mathbf{C}}$ is determined by any three distinct points it contains. Thus, by III.7, in order to find the image under a given linear fractional transformation of a given clircle, it suffices to find the images of three points on that clircle. One can obtain a good geometric picture of any linear-fractional transformation on the basis of this property and two others: a linear-fractional transformation is conformal, and, being a continuous function, it maps connected sets onto connected sets. This remark will be illustrated with the function $\phi(z) = \dfrac{z+i}{z-i}$. (The reader is advised to supplement the following discussion with appropriate sketches.)

First, we have immediately $\phi(-i) = 0$, $\phi(0) = -1$, $\phi(i) = \infty$. The clircle determined by $-i$, 0, i is the extended imaginary axis, and the clircle determined by 0, -1, ∞ is the extended real axis. Hence ϕ maps the extended imaginary axis onto the extended real axis.

The extended imaginary axis separates $\overline{\mathbf{C}}$ into two half-planes, the right half-plane, $\operatorname{Re} z > 0$, which we denote by P_+, and the left half-plane, $\operatorname{Re} z < 0$, which we denote by P_-. Similarly, the extended real axis separates $\overline{\mathbf{C}}$ into the upper half-plane, $\operatorname{Im} z > 0$, which we denote by H_+, and the lower half-plane, $\operatorname{Im} z < 0$, which we denote by H_-. Because ϕ is a bijection of $\overline{\mathbf{C}}$ sending the extended imaginary axis onto the extended real axis, it must map $P_+ \cup P_-$ onto $H_+ \cup H_-$:

$$\phi(P_+) \cup \phi(P_-) = H_+ \cup H_-.$$

Because the sets P_+ and P_- are connected, so are their images, $\phi(P_+)$ and $\phi(P_-)$. The set $\phi(P_+)$ thus cannot intersect both H_+ and H_-, for if it did, the decomposition

$$\phi(P_+) = [\phi(P_+) \cap H_+] \cup [\phi(P_+) \cap H_-]$$

would be a separation of $\phi(P_+)$ into the union of two nonempty sets each disjoint from the closure of the other. Hence, either $\phi(P_+) \subset H_+$ or $\phi(P_+) \subset H_-$. To determine which inclusion is correct, it suffices to note that $\phi(1) = \frac{1+i}{1-i} = i$, which is in H_+. Thus $\phi(P_+) \subset H_+$. By the same reasoning, either $\phi(P_-) \subset H_+$ or $\phi(P_-) \subset H_-$. Because ϕ maps $\overline{\mathbf{C}}$ onto $\overline{\mathbf{C}}$, it cannot map both P_+ and P_- into H_+, so it must be that $\phi(P_-) \subset H_-$. Moreover, the inclusions $\phi(P_+) \subset H_+$ and $\phi(P_-) \subset H_-$ cannot be proper, again because $\phi(\overline{\mathbf{C}}) = \overline{\mathbf{C}}$. Hence $\phi(P_+) = H_+$ and $\phi(P_-) = H_-$.

The unit circle, $|z| = 1$, is sent by ϕ onto a clircle that contains 0 $(= \phi(-i))$ and ∞ $(= \phi(i))$. By the conformality of ϕ, the image of the unit circle intersects the real axis orthogonally at the origin $(= \phi(-i))$. Hence ϕ sends the unit circle onto the extended imaginary axis.

Let D denote the interior of the unit circle, the set $|z| < 1$, and E the extended exterior, the set $1 < |z| \leq \infty$. A repetition of the reasoning above shows that either $\phi(D) = P_+$ or $\phi(D) = P_-$. Since $\phi(0) = -1$, the second equality is correct. Reasoning again as above, we find that $\phi(E) = P_+$.

The image under ϕ of the extended real axis is a clircle that contains the points i $(= \phi(1))$, -1 $(= \phi(0))$ and 1 $(= \phi(\infty))$. Hence ϕ maps the extended real axis onto the unit circle. Another repetition of the reasoning above shows that $\phi(H_+) = E$ and $\phi(H_-) = D$.

We can enhance the picture already obtained by determining the image under ϕ of the coordinate grid (the families of vertical and horizontal lines). An extended vertical line other than the imaginary axis must be sent by ϕ to a circle containing the point 1 $(= \phi(\infty))$ and orthogonal to the unit circle (the image of the extended real axis). The family of extended vertical lines is thus sent to the family of circles tangent to the real axis at the point 1 (plus the real axis itself). By the conformality of ϕ, the family of extended horizontal lines must be sent to the family of clircles orthogonal to the preceding ones. Thus, extended horizontal lines are sent to circles through the point 1 orthogonal to the real axis, except for the extended horizontal line through the point i, which is sent to the extended vertical line through the point 1.

Exercise III.9.1. Find the image of the half-plane $\mathrm{Re}\, z > 0$ under the linear-fractional transformation that maps $0, i, -i$ to $1, -1, 0$, respectively.

Exercise III.9.2. Find the images of the disk $|z| < 1$ and the half-plane $\mathrm{Re}\, z > 0$ under the linear-fractional transformation that maps ∞ to 1 and has i and $-i$ as fixed points.

Exercise III.9.3. Prove that a linear-fractional transformation maps the half-plane $\operatorname{Im} z > 0$ onto itself if and only if it is induced by a matrix with real entries whose determinant is 1.

Exercise* III.9.4. Prove that the linear-fractional transformations mapping the disk $|z| < 1$ onto itself are those of the form $\phi(z) = \dfrac{\lambda(z - z_0)}{\bar{z}_0 z - 1}$, where $|z_0| < 1$ and $|\lambda| = 1$.

Exercise III.9.5. Prove that the linear-fractional transformations mapping the disk $|z| < 1$ onto itself are those induced by matrices of the form

$$\begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix}$$

with $|a|^2 - |b|^2 = 1$.

Exercise III.9.6. Find the image under the linear-fractional transformation $\phi(z) = \dfrac{z-1}{z+1}$ of the intersection of the disk $|z| < 1$ with the half-plane $\operatorname{Im} z > 0$.

Exercise III.9.7. For the function

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

(defined to equal 1 at $z = \infty$ and ∞ at $z = 1$), find the images of the following sets:

(a) The extended real axis.

(b) The extended imaginary axis.

(c) The half-plane $\operatorname{Re} z > 0$.

Exercise III.9.8. Find all linear-fractional transformations that fix the points 1 and -1. Is the group of those transformations abelian? Is it isomorphic to any group with which you are familiar?

Elementary Functions

By the elementary functions of calculus one ordinarily means, in addition to the rational functions, those on the following list:

1. The exponential function, e^x.

2. The logarithm function, $\ln x$.

3. The trigonometric and inverse trigonometric functions, $\sin x$, $\arcsin x$, etc.

4. All functions obtainable in finitely many steps from those already mentioned by means of the operations of addition, subtraction, multiplication, division, and composition. (An example is $\sqrt{1-x^2} = e^{\frac{1}{2}\ln(1-x^2)}$.)

We have already encountered complex rational functions; they are holomorphic and they generalize real rational functions in an obvious way. We shall now see that the other basic elementary functions also have complex holomorphic versions.

IV.1. Definition of e^z

To motivate the definition we follow L. Euler (1707–1783) and perform a simple manipulation with power series, disregarding questions of convergence. One expects that, if there is a natural way of defining e^z, it should maintain the law of exponents: $e^{z_1+z_2} = e^{z_1}e^{z_2}$. Thus, writing $z = x + iy$ as usual, we should have $e^z = e^x e^{iy}$. As we know what e^x means from real-variable theory, it only remains to make sense of e^{iy}. Euler's idea was to write out e^{iy} as a power series, and to fiddle a bit with that series. Since $e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$, we expect to have $e^{iy} = \sum_{n=0}^{\infty}\frac{(iy)^n}{n!}$. Summing separately the even and odd terms of the last series, we obtain

$$\begin{aligned} e^{iy} &= \sum_{k=0}^{\infty} \frac{i^{2k} y^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{i^{2k+1} y^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k y^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} \frac{(-1)^k y^{2k+1}}{(2k+1)!}. \end{aligned}$$

We recognize that the right side expresses in power series form the function $\cos y + i \sin y$.

The considerations above suggest how we should define e^z, namely, by setting

$$e^z = e^x(\cos y + i \sin y).$$

We adopt this as our definition. (The formal manipulations with power series that led to the definition can be made rigorous, as will be clear in the next chapter, where power series are studied systematically.)

As in real-variable theory, we write $\exp z$ instead of e^z when convenient.

Exercise IV.1.1. Find the real and imaginary parts of the function $\exp(e^z)$.

IV.2. Law of Exponents

The expression defining e^z is in polar form: we have $|e^z| = e^x$ and $\arg e^z = y$. The law of exponents, $e^{z_1}e^{z_2} = e^{z_1+z_2}$, follows immediately from the rule for multiplying complex numbers in polar form.

Note that, having defined the exponential function, we can rewrite the polar form, $r(\cos\theta + i\sin\theta)$, of a complex number, as $re^{i\theta}$.

IV.3. e^z is holomorphic

We prove that e^z is holomorphic by verifying that its real and imaginary parts satisfy the Cauchy-Riemann equations. This is a simple calculation. We have $\operatorname{Re} e^z = e^x \cos y$ and $\operatorname{Im} e^z = e^x \sin y$. Since

$$\begin{aligned} &\frac{\partial}{\partial x}(e^x \cos y) = e^x \cos y, &\quad& \frac{\partial}{\partial y}(e^x \cos y) = -e^x \sin y \\ &\frac{\partial}{\partial x}(e^x \sin y) = e^x \sin y, &\quad& \frac{\partial}{\partial y}(e^x \sin y) = e^x \cos y, \end{aligned}$$

the Cauchy-Riemann equations hold. By II.7 we can conclude that e^z is holomorphic, and from II.5 we see that its derivative is e^z.

Exercise* IV.3.1. Suppose the function f is holomorphic in $\mathbf{C}$ and satisfies $f' = f$. Prove that f is a constant times e^z. (Suggestion: Consider the function $e^{-z}f(z)$.)

Exercise IV.3.2. What is the most general holomorphic function f in $\mathbf{C}$ that satisfies $f' = cf$, where c is a constant?

IV.4. Periodicity

Because the functions $\cos y$ and $\sin y$ are periodic with period 2π, we see that the function e^z is periodic with period $2\pi i$: $e^{z+2\pi i} = e^z$. In particular, since $e^0 = 1$, we have $e^{2\pi i} = 1$, a remarkable identity discovered by Euler. It is a favorite of mathomaniacs because it contains in the same short expression the numbers 1, 2, e, π, and i, arguably the five most important constants in mathematics.

Exercise IV.4.1. Prove that if a and b are complex numbers such that $e^a = e^{a+b}$, then b is an integer multiple of $2\pi i$.

IV.5. e^z as a map

Since $|e^z| = e^x$, the map $z \mapsto e^z$ sends the vertical line $x = x_0$ onto the circle with center 0 and radius e^{x_0}. As a point moves along the vertical line in the upward direction, its image point on the circle moves in the counterclockwise direction, making one complete circuit each time its pre-image traverses an interval of length 2π.

Since $\arg e^z = y$, the map $z \mapsto e^z$ maps the horizontal line $y = y_0$ onto the open ray emanating from the origin that makes the angle y_0 with the positive real axis. As a point moves toward the right on the horizontal line, its image point on the ray moves away from the origin.

Thus, the map $z \mapsto e^z$ sends the coordinate grid for rectangular coordinates to the coordinate grid for polar coordinates. Its range is $\mathbf{C}\backslash\{0\}$, and every point in the range has infinitely many pre-images, any two of which differ by an integer multiple of $2\pi i$. The restriction of the map $z \mapsto e^z$ to an open horizontal strip of width at most 2π is one-to-one, and the map sends the strip to the interior of an angle with vertex at the origin. For example, the strip

$$\frac{-\pi}{2} < y < \frac{\pi}{2}$$

is sent onto the right half-plane.

Exercise IV.5.1. Describe the images under the map $z \mapsto e^z$ of the line $y = x$ and of the strip

$$x - \frac{\pi}{2} < y < x + \frac{\pi}{2}.$$

Exercise IV.5.2. Describe the curves $|f|$ = constant and $\arg f$ = constant for the function

$$f(z) = \exp(z^2).$$

Exercise IV.5.3. Repeat Exercise IV.5.2 for the function

$$f(z) = \exp\left(\frac{z+1}{z-1}\right).$$

IV.6. Hyperbolic functions

The hyperbolic functions are defined in terms of the exponential function just as in the theory of functions of a real variable:

$$\begin{aligned} \cosh z &= \frac{e^z + e^{-z}}{2}, & \sinh z &= \frac{e^z - e^{-z}}{2}, \\ \tanh z &= \frac{\sinh z}{\cosh z}, & \coth z &= \frac{\cosh z}{\sinh z}, \\ \operatorname{sech} z &= \frac{1}{\cosh z}, & \operatorname{csch} z &= \frac{1}{\sinh z}. \end{aligned}$$

Because the function e^z is holomorphic, so are the hyperbolic functions at every point where they are defined, which is everywhere for the first two, and everywhere the denominator in the defining expression does not vanish for the other four. The expressions for the derivatives of the hyperbolic functions and various identities involving the functions are easily obtained from properties of the exponential function (Exercises IV.6.1 and IV.6.2 below).

Exercise IV.6.1. Derive the expressions for the derivatives of the hyperbolic functions.

Exercise IV.6.2. Derive the identities

$$\cosh(z_1 + z_2) = \cosh z_1 \cosh z_2 + \sinh z_1 \sinh z_2,$$
$$\sinh(z_1 + z_2) = \sinh z_1 \cosh z_2 + \cosh z_1 \sinh z_2.$$

Exercise IV.6.3. Suppose the holomorphic function f in $\mathbf{C}$ has a holomorphic derivative and satisfies the differential equation $f'' = f$. Prove that f has the form $f(z) = a \cosh z + b \sinh z$, where a and b are constants.

IV.7. Zeros of cosh$z$ and sinh$z$.

The zeros of the function $\sinh z$ are the solutions for z of the equation $e^z = e^{-z}$, which can be rewritten as $e^{2z} = 1$. Since $e^{2z} = 1$ only when $2z$ is an integer multiple of $2\pi i$, we conclude that the zeros of $\sinh z$ are the numbers $n\pi i$ with n an integer.

Similarly, the zeros of $\cosh z$ are the solutions for z of the equation $e^z = -e^{-z}$, which can be written as $e^{2z} = -1$, or as $e^{2z-\pi i} = 1$ (since $e^{\pi i} = -1$). It follows that the zeros of $\cosh z$ are the numbers $\left(n + \frac{1}{2}\right)\pi i$ with n an integer.

IV.8. Trigonometric functions

If y is real then we have $\cos y = \dfrac{e^{iy} + e^{-iy}}{2}$ and $\sin y = \dfrac{e^{iy} - e^{-iy}}{2i}$. We define the functions $\cos z$ and $\sin z$ by replacing y by z in these expressions. Our definitions are thus

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

The other trigonometric functions are defined in terms of cosine and sine just as in real variable theory:

$$\tan z = \frac{\sin z}{\cos z}, \qquad \cot z = \frac{\cos z}{\sin z},$$
$$\sec z = \frac{1}{\cos z}, \qquad \csc z = \frac{1}{\sin z}.$$

These functions are holomorphic, the first two everywhere, the other four everywhere their denominators are nonzero.

In previous encounters with trigonometric and hyperbolic functions, the reader has perhaps taken note of the formal similarities between the two function classes. The addition formula for the hyperbolic sine, for example, is the same as that for the sine, and the addition formula for the hyperbolic cosine differs from that for the cosine only in a minus sign. The complex viewpoint explains these similarities, because from the basic definitions we have the relations

$$\cos z = \cosh iz, \qquad \sin z = \frac{\sinh iz}{i}.$$

Roughly speaking, one obtains the trigonometric functions by composing the hyperbolic functions with a rotation of the complex plane through 90 degrees.

The basic properties of the trigonometric functions are easily deduced from those of the exponential function and the hyperbolic functions. For example, from what we already know about the latter functions, we see that the functions $\cos z$ and $\sin z$ are periodic with period 2π, that the zeros of $\sin z$ are the numbers $n\pi$ with n an integer, and that the zeros of $\cos z$ are the numbers $\left(n + \frac{1}{2}\right)\pi$ with n an integer. In particular, in extending the functions sine and cosine from $\mathbf{R}$ to $\mathbf{C}$ we have not introduced additional zeros.

Further properties are left for the reader to establish.

Exercise IV.8.1. Find all the roots of the equation $\cos z = 2$.

Exercise IV.8.2. Establish the identities

$$\cos z = \cos x \cosh y - i \sin x \sinh y,$$

$$\sin z = \sin x \cosh y + i \cos x \sinh y,$$
$$|\cos z|^2 = \cos^2 x + \sinh^2 y,$$
$$|\sin z|^2 = \sin^2 x + \sinh^2 y.$$

Exercise IV.8.3. Describe the images of the lines Re z = constant and Im z = constant under the map $z \mapsto \cos z$.

IV.9. Logarithms

If a is a complex number, then by a logarithm of a one means any complex number b such that $e^b = a$. If $a = 0$ there is no such b, but otherwise there are infinitely many, any two of which differ by an integer multiple of $2\pi i$. For example, the logarithms of the number 1 are the numbers $2n\pi i$ with n an integer, the logarithms of the number -1 are the numbers $(2n+1)\pi i$ with n an integer, and the logarithms of the number i are the numbers $\left(2n + \frac{1}{2}\right)\pi i$ with n an integer.

Since $|e^b| = e^{\operatorname{Re} b}$ and arg e^b = Im b, the number b is a logarithm of the nonzero number a if and only if $e^{\operatorname{Re} b} = |a|$ and Im b = arg a. Thus, letting ln denote the natural logarithm function of calculus, we see that the logarithms of a are the numbers $\ln|a| + i$ arg a. If b is a logarithm of a, we shall write $b = \log a$. The particular logarithm of a whose imaginary part is in the interval $(-\pi, \pi]$, corresponding to the principal value of arg a, will be denoted by Log a; i.e., Log $a = \ln|a| + i \operatorname{Arg} a$.

Exercise IV.9.1. Find all values of $\cosh(\log 2)$.

Exercise IV.9.2. Find all values of $\log(\log i)$.

Exercise IV.9.3. In what sense is it true that $\log a_1 a_2 = \log a_1 + \log a_2$ for complex numbers a_1 and a_2?

IV.10. Branches of $\arg z$ and $\log z$.

Let G be an open connected subset of $\mathbf{C}$ not containing the origin. By a branch of arg z in G is meant a continuous function α in G such that, for each z in G, the value $\alpha(z)$ is a value of arg z. By a branch of log z in G is meant a continuous function l in G such that, for each z in G, the value $l(z)$ is a logarithm of z. For a simple example, we can take for G the region one obtains by removing from $\mathbf{C}$ the interval $(-\infty, 0]$ on the real axis; in this region the functions $\alpha(z)$ = Arg z and $l(z)$ = Log z are branches of arg z and log z, respectively.

As we shall see, depending on the shape of G, there may or may not exist branches of arg z and log z in G. However, if there is a branch α of arg z in G then there is a branch of log z, namely, the function $l(z) = \ln|z| + i\alpha(z)$.

Similarly, if there is a branch l of $\log z$ in G then there is a branch of $\arg z$, namely, the function $\alpha(z) = \operatorname{Im} l(z)$.

If α is a branch of $\arg z$ in G then $\alpha + 2\pi n$ is another branch for any integer n. On the other hand, if α_1 and α_2 are two branches of $\arg z$ in G, then the function $\dfrac{\alpha_1 - \alpha_2}{2\pi}$ is continuous in G and takes only integer values, so it must be constant (because by assumption G is connected). Thus, if there is a branch of $\arg z$ in G, then any two branches differ by a constant integer multiple of 2π. Similarly, if there is a branch l of $\log z$ in G, then $l + 2\pi n i$ is another branch for every integer n, and every branch has this form.

An example of a region that carries no branch of $\arg z$ is given in the exercise at the end of the section. On the other hand, if G is any open disk that excludes the origin, then G does carry such a branch. In fact, suppose θ_0 is a particular value of the argument of the center of the disk. Then the disk is contained in the half-plane $\left\{re^{i\theta} : r > 0, \theta_0 - \dfrac{\pi}{2} < \theta < \theta_0 + \dfrac{\pi}{2}\right\}$, and we can obtain the desired branch α by letting $\alpha(z)$, for z in G, equal the value of $\arg z$ in the interval $\left(\theta_0 - \dfrac{\pi}{2},\ \theta_0 + \dfrac{\pi}{2}\right)$.

Exercise* IV.10.1. Prove that there is no branch of $\arg z$ in the region $0 < |z| < 1$.

IV.11. $\log z$ as a holomorphic function

If l is a branch of $\log z$ in the open connected set G, then l is holomorphic, and $l'(z) = \dfrac{1}{z}$.

We shall verify that l is holomorphic by checking that l satisfies the Cauchy-Riemann equations in polar coordinates, whose derivation is Exercise II.6.3. Let $u = \operatorname{Re} l$ and $v = \operatorname{Im} l$. The Cauchy-Riemann equations in polar coordinates are

$$r\frac{\partial u}{\partial r} = \frac{\partial v}{\partial \theta}, \qquad \frac{\partial u}{\partial \theta} = -r\frac{\partial v}{\partial r}.$$

Since l is a branch of $\log z$ we have $u(z) = \ln|z|$ and $v(z) = \arg z$, from which it is immediate that

$$\frac{\partial u}{\partial r} = \frac{1}{r}, \qquad \frac{\partial u}{\partial \theta} = 0 = \frac{\partial v}{\partial r}, \qquad \frac{\partial v}{\partial \theta} = 1.$$

The Cauchy-Riemann equations are thus satisfied, and we can conclude by II.7 that l is holomorphic. Moreover, by II.6 we have $l' = \dfrac{\partial u}{\partial x} + i\dfrac{\partial v}{\partial x}$.

Using the chain rule, we obtain

$$l'(z) = \frac{x - iy}{r^2} = \frac{1}{x + iy} = \frac{1}{z}.$$

IV.12. Logarithms of holomorphic functions

Let G be an open connected subset of $\mathbf{C}$, and let f be a holomorphic function in G that does not assume the value 0. By a branch of $\log f$ in G one means a continuous function g in G such that $f = e^g$, in other words, such that, for each z in G, the value $g(z)$ is a logarithm of $f(z)$. The special case of the function $f(z) = z$ is the one discussed in the three preceding sections. Here in general, just as in that special case, one sees that if there is a branch of $\log f$ in G, then one obtains the most general such branch by adding a constant integer multiple of $2\pi i$ to any particular one.

If it happens that the range of f is contained in a region in which there is a branch l of $\log z$, then the composite function $l \circ f$ is a branch of $\log f$; moreover, being a composite of two holomorphic functions it is holomorphic, and by the chain rule its derivative equals $\frac{f'}{f}$.

Whether or not there is a branch of $\log f$ in G, such branches exist locally, in the following sense. Fix a point z_0 in G. Since by assumption $f(z_0) \neq 0$, there is by the continuity of f an open disk D with center z_0 such that $f(D)$ is contained in an open disk D' that excludes the origin. Then, as seen in Section IV.10, there is a branch l of $\log z$ in D', and by composing l with f we obtain a branch of $\log f$ in D. In case there happens to be a branch g of $\log f$ in all of G, the functions g and $l \circ f$ differ in D by a constant integer multiple of $2\pi i$. In fact, we can arrange to have $g = l \circ f$ by choosing for l the particular branch of $\log z$ in D' satisfying $l\big(f(z_0)\big) = g(z_0)$. Since $l \circ f$ is a holomorphic function with derivative $\frac{f'}{f}$, we can draw the following conclusion: *If there is a branch of* $\log f$ *in G, then any such branch is a holomorphic function whose derivative is* $\frac{f'}{f}$. Whether or not there is a branch of $\log f$ in G, one refers to the function $\frac{f'}{f}$ as the logarithmic derivative of f.

Although, as just seen, a branch of $\log f$ in G, if it exists, is locally the composite of f with a branch of $\log z$, it may not be globally such a composite. A simple example is provided by the function $f(z) = e^z$ in $\mathbf{C}$. The function $g(z) = z$ is a branch of $\log f$, yet the range of f is $\mathbf{C}\backslash\{0\}$, in which there is no branch of $\log z$ (according to Exercise IV.10.1). Whether, for a particular f and G, there exists a branch of $\log f$ in G, is a subtle question which we shall only be able to deal with adequately much later, in Chapter X.

Exercise IV.12.1. Prove that the logarithmic derivative of the product of two holomorphic functions equals the sum of their logarithmic derivatives.

IV.13. Roots

Let G be a connected open subset of $\mathbf{C}$ and let f be a holomorphic function in G that does not assume the value 0. Let n be a positive integer. By a branch of $f^{\frac{1}{n}}$ in G one means a continuous function h in G such that, for each z in G, the value $h(z)$ is an n-th root of $f(z)$. (The presumption that f omits the value 0 is explained by Exercise IV.13.1 below.) The discussion of branches of $f^{\frac{1}{n}}$ parallels that of branches of $\log f$ and will only be briefly sketched. The basic facts are as follows. If there is a branch of $f^{\frac{1}{n}}$ in G then there are precisely n such branches, each obtainable from any particular one through multiplication by an n-th root of 1. If there is a branch g of $\log f$ in G, then the function $e^{\frac{g}{n}}$ is a branch of $f^{\frac{1}{n}}$ in G. Any branch of $f^{\frac{1}{n}}$ can be obtained locally in the preceding way from a local branch of $\log f$. If h is a branch of $f^{\frac{1}{n}}$ then h is holomorphic and $\dfrac{h'}{h} = \dfrac{f'}{nf}$.

Exercise* IV.13.1. Prove that for $n > 1$ there is no branch of $z^{\frac{1}{n}}$ in the region $0 < |z| < 1$.

Exercise IV.13.2. Prove that if h is a branch of $f^{\frac{1}{n}}$ then h is holomorphic and $\dfrac{h'}{h} = \dfrac{f'}{nf}$.

Exercise IV.13.3. Let G be the open set one obtains by removing from $\mathbf{C}$ the interval $[-1, 1]$ on the real axis. Prove that there is a branch of the function $\sqrt{\dfrac{z+1}{z-1}}$ in G. (Suggestion: What is the image of G under the map $z \mapsto \dfrac{z+1}{z-1}$?)

Exercise IV.13.4. Let G be as in Exercise IV.13.3. Prove that there is a branch of the function $\sqrt{z^2 - 1}$ in G.

IV.14. Inverses of holomorphic functions

The use of branches gives a way of dealing with inverses of functions that are not one-to-one. This device appears already in the theory of functions of a real variable. For example, the function $f(x) = x^2$ on $\mathbf{R}$ has range $[0, \infty)$, and each point in its range, except for the origin, has two pre-images. The "inverse function" has two branches, namely, the function $g(x) = \sqrt{x}$ (the positive square root), and the function $-g$. The function g is the bona fide inverse of the restriction of f to the interval $[0, \infty)$, while $-g$ is the inverse of the restriction of f to the interval $(-\infty, 0]$; both of those restrictions of f are one-to-one functions. Similarly, the sine function on $\mathbf{R}$ has range [-1,1], and each point in its range has infinitely many pre-images. The so-called inverse sine function, the function arcsin, is the inverse of the restriction of the sine function to the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$; it is only one of infinitely many branches of the inverse of the sine function.

Suppose f is a holomorphic function in the open connected subset G of $\mathbf{C}$. If f is one-to-one, or, in the parlance of complex analysis, "univalent," then the inverse of f is the function g defined in the region $f(G)$ by the rule $g(f(z)) = z$. It turns out, as we shall see later (in Chapter X), that if f is univalent then $f(G)$ is an open set; in fact, that much is true under the weaker assumption that f is nonconstant. Moreover, its derivative, f', is never 0, and its inverse function is continuous. Granted these facts, one easily sees, by the same reasoning one uses in calculus, that the inverse function is holomorphic, its derivative at the point $f(z)$ of $f(G)$ being equal to $\frac{1}{f'(z)}$ (Exercise IV.14.1 below).

If the holomorphic function f, defined in the open connected set G, is not univalent, then it does not possess an inverse function in the strict sense of the term. Nevertheless, it is still possible to speak of branches of the inverse of f. All one needs is an open connected subset H of G on which f is univalent; the inverse of the restriction of f to H is then called a branch of the inverse of f. Starting with the function $f(z) = e^z$ one obtains in this way what we called the branches of $\log z$; starting with the function $f(z) = z^n$ one obtains the branches of $z^{\frac{1}{n}}$.

Exercise* IV.14.1. Let f be a univalent holomorphic function in the open connected set G, and let g be the inverse function. Assume that $f(G)$ is open, that g is continuous, and that f' is never 0. Prove g is holomorphic.

Exercise IV.14.2. Prove that a branch of the inverse of a holomorphic function is always univalent.

IV.15. Inverse trigonometric functions

The preceding discussion, applied to any of the trigonometric functions, explains what one means by the branches of the corresponding inverse function. Because of the relation between the trigonometric functions and the exponential function, those branches can be related to the logarithm function.

To illustrate we consider the cosine function; the branches of its inverse are referred to as the branches of $\arccos z$. Let us fix a point z in the domain of such a branch and ask what possible values the branch can have at z. If w is such a value, then w is a solution of the equation $\cos w = z$; in other words, $\dfrac{e^{iw}+e^{-iw}}{2} = z$, which can be rewritten as

$$e^{2iw} - 2ze^{iw} + 1 = 0.$$

This is a quadratic equation in e^{iw}. The quadratic formula gives us two roots (or one repeated one), namely $e^{iw} = z + \sqrt{z^2-1}$ (recall that the square root has two possible values if $z \neq \pm 1$). The possible values of w are therefore expressed by $w = \dfrac{1}{i}\log\left(z+\sqrt{z^2-1}\right)$. One sees in this way that the branches of $\arccos z$ are just the branches of $\dfrac{1}{i}\log\left(z+\sqrt{z^2-1}\right)$.

Exercise IV.15.1. Find a relation between the branches of $\arctan z$ and the logarithm function.

IV.16. Powers

If a and c are complex numbers, with $a \neq 0$, then by the values of a^c one means the values of $e^{c\log a}$. For example, the values of 1^i are the numbers $e^{i(2\pi n i)}$, in other words, the numbers $e^{-2\pi n}$, with $n = 0, \pm 1, \pm 2, \cdots$. If c is an integer or the reciprocal of an integer, the preceding definition reduces to the usual one.

On the basis of the preceding definition one can give a meaning, for any complex number c, to the branches of z^c, or, more generally, to the branches of f^c for any holomorphic function f that omits the value 0. As this involves no ideas not encountered earlier, the details will not be spelled out.

Exercise IV.16.1. Find all the values of $(1+i)^i$.

Exercise IV.16.2. Let a be a complex number of unit modulus and c an irrational real number. Prove that the values of a^c form a dense subset of the unit circle.

Exercise IV.16.3. Prove that if f is a branch of z^c in an open set not containing 0, then f is holomorphic and f' is a branch of cz^{c-1}.

IV.17. Analytic continuation and Riemann surfaces

The notion of a branch has enabled us to make sense of so-called multiple-valued functions like $\log z$ and $\sqrt{z}$. What we have done so far in this direction, although it will be adequate for our purposes, is really only the first step in a more profound treatment of multiple-valued functions. This more profound treatment is generally not included in an introductory course; it relies on results we shall only develop later on. Its main ideas will be very vaguely sketched, with emphasis on the simple example $\sqrt{z}$, in this concluding section of Chapter IV.

A key observation is that the different branches of $\sqrt{z}$ are related to one another through a process called analytic continuation. To illustrate, consider the four open disks D_0, D_1, D_2, D_3, each of unit radius, with respective centers 1, i, -1, $-i$. In D_0 there are two branches of $\sqrt{z}$, which we call f and g, one of which, say f, takes the value 1 at $z = 1$, and the other the value -1 at $z = 1$. In D_1 there are also two branches of $\sqrt{z}$, and one of them, which we call f_1, agrees with f in the overlap $D_0 \cap D_1$. The reader will easily check that $f_1(i) = e^{\frac{\pi i}{4}}$. Similarly, in D_2 there are two branches of $\sqrt{z}$, one of which, say f_2, agrees with f_1 in the overlap $D_1 \cap D_2$. One then has $f_2(-1) = i$. We continue in this way for two more steps. The branch of $\sqrt{z}$ in D_3 that agrees with f_2 in $D_2 \cap D_3$ will be called f_3; it satisfies $f_3(-i) = e^{\frac{3\pi i}{4}}$. Finally, the branch of $\sqrt{z}$ in D_0 that agrees with f_3 in $D_3 \cap D_0$ is g (not f). The two branches f and g of $\sqrt{z}$ that live in D_0 are thus linked by the successive branches f_1, f_2, f_3; one says that g arises from f by analytic continuation. A precise definition of the latter term will not be attempted—it will only be mentioned that any two branches of $\sqrt{z}$ are linked in this way, and that through analytic continuation of a branch of $\sqrt{z}$ only other branches of $\sqrt{z}$ can arise.

It thus appears that the totality of the branches of $\sqrt{z}$ form some kind of organic whole, and it seems sensible to conceive of $\sqrt{z}$ as the collection of all of its branches. Except for certain technical modifications, this is what is actually done. When $\sqrt{z}$ is interpreted this way, it acquires a natural topological structure that makes it into a surface. That surface can be visualized as follows.

We consider two copies of $\mathbf{C}$, each slit along the negative real axis; we imagine them as stacked one above the other in real three-dimensional space. Let A denote the one on top and B the bottom one. (It might help to associate A with one of the branches of $\sqrt{z}$ in $\mathbf{C}\backslash(-\infty, 0]$ and B with the other branch.) We join A and B, making a single surface, by identifying the top edge of the slit in A with the bottom edge of the slit in B, and vice versa (see Figure 3). As there are to be no additional identifications, the resulting surface cannot actually be embedded in real three-dimensional space, but by cheating a little we can imagine it as so embedded. Set-theoretically the construction poses no difficulty, although we shall not worry here about the technical details. Let S denote the resulting surface. From the way we constructed S, out of two copies of $\mathbf{C}$, we obtain a natural map from S to $\mathbf{C}$. The map is two-to-one but locally one-to-one, so by using it we can introduce local complex coordinates in S, thereby obtaining a way to define what it means for a complex-valued function on S to be holomorphic, or for a function of $\mathbf{C}$ into S to be holomorphic. (There might seem to be trouble at the origin, but the construction is performed in such a way that the origin is deleted from S.) We can reinterpret the function z^2 in a natural way as a univalent function from $\mathbf{C}\backslash\{0\}$ onto S; that function then has a well-defined inverse, a function mapping S to $\mathbf{C}$. In this way we can think of $\sqrt{z}$, not as a "multiple-valued function" in $\mathbf{C}$, but as a bona fide (single-valued) function on the surface S.

A similar construction is possible for other multiple-valued functions. For example, to form a surface that carries $\log z$ as a single-valued function, one uses countably many copies of $\mathbf{C}$, each slit along the negative real axis, indexed by all of the integers, and imagined as stacked one above the other in three-dimensional space. One makes a single surface by identifying the upper edge of the slit in each plane with the lower edge of the slit in the plane immediately above it. As in the preceding example, there is a natural way of introducing local complex coordinates on this surface.

The surfaces associated with the functions $\sqrt{z}$ and $\log z$ are examples of what are called Riemann surfaces. One can associate such a surface with any multiple-valued holomorphic function. While for simple cases like $\sqrt{z}$ and $\log z$ the surface can be described concretely, as indicated above, in the general case it is constructed in a natural way from the branches of the function. The surfaces serve to make mathematically rigorous the intuitive notion of multiple-valued function; at the same time they provide domains on which their associated multiple-valued functions become single-valued.

In general, a Riemann surface is a surface equipped in a consistent way with local complex coordinates. Such a surface need not arise from a multiple-valued function. The simplest example of one that does not is $\overline{\mathbf{C}}$, the extended complex plane (or Riemann sphere). To make $\overline{\mathbf{C}}$ into a Riemann surface one needs only to introduce complex coordinates in a neighborhood of ∞, which one does by means of the map $z \mapsto \dfrac{1}{z}$. It was Riemann's brilliant insight that Riemann surfaces, and not merely the complex plane itself, form the natural setting for complex analysis. The study of Riemann surfaces is a major activity in modern complex function theory.

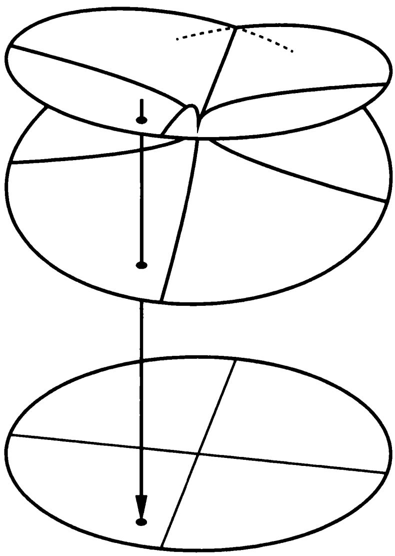

Figure 3. A depiction of the Riemann surface of $\sqrt{z}$, with the natural map from the surface to $\mathbf{C}$.

Chapter V

Power Series

Power series play a central role in complex function theory because, as we shall see in Chapter VII, every holomorphic function can be represented locally by such series. The basic theory of complex power series, to be developed in this chapter, is largely indistinguishable from its real counterpart.

V.1. Infinite Series

Given a sequence $c_0,\ c_1,\ c_2, \ldots$ of complex numbers, one says that the infinite series $\sum_{n=0}^{\infty} c_n$ converges if $\lim\limits_{N\to\infty} \sum_{n=0}^{N} c_n$ exists and is finite. The number $s = \lim\limits_{N\to\infty} \sum_{n=0}^{N} c_n$ is then called the sum of the series, written $s = \sum_{n=0}^{\infty} c_n$. The number c_n is called the n-th term and the number $s_N = \sum_{n=0}^{N} c_n$ the N-th partial sum of the series. A series that does not converge is said to diverge.

V.2. Necessary Condition for Convergence

If the series $\sum_{n=0}^{\infty} c_n$ *converges then* $\lim\limits_{n\to\infty} c_n = 0$.

The proof is the same as for real series. Namely, assume the series converges, let s be its sum, and s_N its N-th partial sum. Then $s_N - s_{N-1} = c_N$, so

$$\begin{aligned}\lim_{N\to\infty} c_N &= \lim_{N\to\infty} (s_N - s_{N-1}) \\ &= \lim_{N\to\infty} s_N - \lim_{N\to\infty} s_{N-1} \\ &= s - s = 0.\end{aligned}$$

(Use has been made of the fact that, if two sequences of complex numbers converge, then the sequence of differences converges to the difference of the limits. This is a corollary of Exercise I.8.3.)

V.3. Geometric Series

If c is a complex number, then the series $\sum_{n=0}^{\infty} c^n$ converges to $\frac{1}{1-c}$ if $|c| < 1$ and diverges if $|c| \geq 1$.

The statement about divergence follows immediately from (V.2). The statement about convergence is a consequence of the identity

$$\sum_{n=0}^{N} c^n = \frac{1-c^{N+1}}{1-c},$$

which holds whenever $c \neq 1$, and which one can easily verify by multiplying both sides by $1-c$ and making cancellations on the left side. If $|c| < 1$ then $|c|^{N+1} \longrightarrow 0$ as $N \longrightarrow \infty$, and the desired conclusion follows.

V.4. Triangle Inequality for Series

If the series $\sum_{n=0}^{\infty} c_n$ converges, then $\left|\sum_{n=0}^{\infty} c_n\right| \leq \sum_{n=0}^{\infty} |c_n|$.

In fact, let s denote the sum of the series and s_N its N-th partial sum. We have

$$|s_N| \leq \sum_{n=0}^{N} |c_n|$$

(by Exercise I.5.1). As $N \longrightarrow \infty$ the left side converges to $|s|$ (by Exercise I.5.2) and the right side converges to $\sum_{n=0}^{\infty} |c_n|$. The desired inequality follows.

V.5. Absolute Convergence

The series $\sum_{n=0}^{\infty} c_n$ is said to converge absolutely if the series $\sum_{n=0}^{\infty} |c_n|$ converges. Just as for real series, the absolute convergence of a complex series implies its convergence. In fact, letting s_N as usual denote the N-th partial sum of the series $\sum_{n=0}^{\infty} c_n$, we have, for $0 < M < N$,

$$|s_N - s_M| = \left|\sum_{n=M+1}^{N} c_n\right| \leq \sum_{n=M+1}^{N} |c_n| \leq \sum_{n=M+1}^{\infty} |c_n|.$$

If the series $\sum_{n=0}^{\infty} |c_n|$ converges, then the quantity on the right side tends to 0 as $M \longrightarrow \infty$, enabling us to conclude that the sequence $(s_N)_{N=1}^{\infty}$ is a Cauchy sequence. Since the metric space $\mathbf{C}$ is complete, the sequence of partial sums converges, as desired.

V.6. Sequences of Functions

Suppose $(g_n)_{n=0}^{\infty}$ is a sequence of complex-valued functions defined in the open subset G of $\mathbf{C}$. The sequence is said to converge on the subset S of G if $\lim_{n\to\infty} g_n(z)$ exists finitely for each z in S. If that happens, and g is the limit function (i.e., $g(z) = \lim_{n\to\infty} g_n(z)$ for each z in S), then one says that the sequence converges uniformly on S provided $\sup\{|g(z)-g_n(z)| : z \in S\}$ tends to 0 as $n \longrightarrow \infty$. Stated in ϵ-language, the sequence converges uniformly to g on S if, for each positive number ϵ, there exists a positive integer n_0 such that $|g(z) - g_n(z)| < \epsilon$ for all $n \geq n_0$ and all z in S. An equivalent condition, which does not mention the limit function, is that the sequence $(g_n)_{n=0}^{\infty}$ be uniformly Cauchy on S, in other words, that for each positive number ϵ, there exist a positive integer n_0 such that $|g_n(z) - g_m(z)| < \epsilon$ for all $m \geq n_0$, all $n \geq n_0$, and all z in S (see Exercise V.6.1 below).

The sequence $(g_n)_{n=0}^{\infty}$ is said to converge locally uniformly in G if each point of G has a neighborhood in which the sequence converges uniformly. A simple example, with G the open unit disk, is provided by the sequence $g_n(z) = z^n$. If $0 < r_0 < 1$ then this sequence converges uniformly to 0 in the disk $|z| < r_0$, and so it converges locally uniformly to 0 in the disk $|z| < 1$. However, the sequence does not converge uniformly in the disk $|z| < 1$.

Exercise* V.6.1. Prove, with the notations above, that the sequence $(g_n)_{n=0}^{\infty}$ converges uniformly on S if and only if it is uniformly Cauchy on S.

Exercise V.6.2. Prove that the sequence $(g_n)_{n=0}^{\infty}$ converges locally uniformly in the open set G if and only if it converges uniformly on each compact subset of G.

V.7. Series of Functions

Consider an infinite series $\sum_{n=0}^{\infty} f_n$, where the terms f_n now are not complex numbers but complex-valued functions, all defined in some open subset G of $\mathbf{C}$. As was the case with series of numbers, we define convergence for a series of functions by referring to the sequence of partial sums. The N-th partial sum of the series $\sum_{n=0}^{\infty} f_n$ is the function $g_N = \sum_{n=0}^{N} f_n$. We say the series converges, or converges uniformly, on the subset S of G if the sequence $(g_N)_{N=0}^{\infty}$ of partial sums does. We say that the series converges locally uniformly in G if the sequence of partial sums does. It is an elementary exercise to show that if the series $\sum_{n=0}^{\infty} |f_n|$ converges locally uniformly in G then so does the series $\sum_{n=0}^{\infty} f_n$ (Exercise V.7.1 below).

As a simple example, consider the geometric series $\sum_{n=0}^{\infty} z^n$. From the discussion in Section (V.3) we know that the series converges in the disk

$|z| < 1$ to $\dfrac{1}{1-z}$, and that

$$\frac{1}{1-z} - \sum_{n=0}^{N} z^n = \frac{z^{N+1}}{1-z}.$$

It follows that the series $\sum_{n=0}^{\infty} z^n$ converges locally uniformly to $\dfrac{1}{1-z}$ in the disk $|z| < 1$.

Exercise* V.7.1. Prove, using the notations above, that the local uniform convergence of the series $\sum_{n=0}^{\infty} |f_n|$ in G implies the local uniform convergence of the series $\sum_{n=0}^{\infty} f_n$ in G.

Exercise* V.7.2. Prove that the series $\sum_{n=0}^{\infty} \left(\dfrac{z-1}{z+1}\right)^n$ converges locally uniformly in the half-plane $\operatorname{Re} z > 0$, and find the sum.

V.8. Power Series

A power series is a series of the form $\sum_{n=0}^{\infty} a_n(z-z_0)^n$, where $z_0,\ a_0,\ a_1,\ a_2, \ldots$ are complex constants. The number z_0 is called the center of the series, and the number a_n is called the n-th coefficient. The series converges at least at $z = z_0$. (Note that, when working with power series, we always use the convention $0^0 = 1$. Thus, when $z = z_0$, the initial term of the series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ has the value a_0.)

Power series can be thought of as "generalized polynomials." As we shall see, they can often be manipulated exactly as if they were polynomials. From this perspective, a polynomial is just a power series in which only finitely many coefficients are nonzero, and for which therefore questions of convergence do not arise.

If a power series converges to a function f in a given region, we shall say that the series represents f in that region. It is important, however, to distinguish the series from the function it represents. For example, as shown at the end of the preceding section, the series $\sum_{n=0}^{\infty} z^n$ represents the function $\dfrac{1}{1-z}$ in the disk $|z| < 1$. However, the series is not "equal" to the function, in a formal sense, even though we can write $\sum_{n=0}^{\infty} z^n = \dfrac{1}{1-z}$ for $|z| < 1$. The same function is represented by other power series; for example, it is represented by the series $\sum_{n=0}^{\infty} 2^{-n-1}(z+1)^n$ in the larger disk $|z+1| < 2$, as the reader will easily verify. A power series is best thought of as a formal sum, uniquely determined once its center and its coefficients have been specified.

Exercise V.8.1. Let k be a positive integer and let z_0 be a point of **C**. Find the power series with center z_0 that represents the function z^k.

V.9. Region of Convergence

The region of convergence of a power series with center z_0 is either the singleton $\{z_0\}$, the entire complex plane, or an open disk with center z_0 plus a subset of the boundary of that disk. In any open disk with center z_0 where it converges, the series converges absolutely and locally uniformly.

Thus, the convergence picture for power series is relatively simple. The preceding proposition describes the situation completely, except for the question of how the series behaves, in the third of the three possibilities above, on the circle forming the boundary between the region of convergence and the region of divergence. Although that is an interesting question which has been the subject of deep studies, it is peripheral to the present development.

In proving the proposition we shall for simplicity treat only the case where $z_0 = 0$. It will be clear that the general case is no different. The proposition is an immediate consequence of the following assertion: *If the power series $\sum_{n=0}^{\infty} a_n z^n$ converges at a point on the circle $|z| = r_0$, where $r_0 > 0$, then it converges absolutely and locally uniformly in the disk $|z| < r_0$.*

We shall prove this assertion by comparing the given series with a geometric series. Suppose the series $\sum_{n=0}^{\infty} a_n z^n$ converges for $z = c$, where $|c| = r_0$. Then $|a_n c^n| \to 0$ as $n \to \infty$ by (V.2), so the number $M = \sup\{|a_n| r_0^n : n = 0, 1, \dots\}$ is finite. Since

$$|a_n z^n| = |a_n| r_0^n \left|\frac{z}{r_0}\right|^n \leq M \left|\frac{z}{r_0}\right|^n,$$

each term of the series $\sum_{n=1}^{\infty} a_n z^n$ is dominated in absolute value by M times the corresponding term of the geometric series $\sum_{n=0}^{\infty} \left|\frac{z}{r_0}\right|^n$. The last series converges locally uniformly in the disk $|z| < r_0$; in fact, its sum in that disk is $\dfrac{1}{(1 - |z/r_0|)}$, and the difference between its sum and its N-th partial sum is $\dfrac{|z/r_0|^{N+1}}{(1 - |z/r_0|)}$ (see Section V.7). Using the Cauchy criterion for uniform convergence (stated in Section V.6, with the proof requested in Exercise V.6.1) we can conclude that the series $\sum_{n=0}^{\infty} |a_n z^n|$ also converges locally uniformly in the disk $|z| < r_0$, which is the desired conclusion (in view of Exercise V.7.1).

V.10. Radius of Convergence

Let $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ be a power series, and let R denote the supremum of $|z-z_0|$ over all numbers z for which the series converges. The number R is called the radius of convergence of the series. By the assertion proved in the preceding section, if $R>0$ then the series converges absolutely and locally uniformly in the disk $|z-z_0|<R$; if $R<\infty$ then the series diverges at each point of the region $|z-z_0|>R$.

Presently we shall obtain a general expression for the radius of convergence of a power series in terms of its coefficients. Its proof involves a comparison with geometric series, as did the proof of V.9. The next exercise makes explicit the simple comparison test involved.

Exercise* V.10.1. Let $\sum_{n=0}^{\infty} a_n z^n$ and $\sum_{n=0}^{\infty} b_n z^n$ be power series with respective radii of convergence R_1 and R_2. Assume that there is a positive constant M such that $|a_n| \le M|b_n|$ for all but finitely many n. Prove that $R_1 \ge R_2$. Use this test to prove that the radius of convergence of the series $\sum_{n=1}^{\infty} n^{-n} z^n$ is ∞.

Exercise V.10.2. Prove that the series $\sum_{n=0}^{\infty} n^n z^n$ has radius of convergence 0.

V.11. Limits Superior

A basic notion from real analysis will now be reviewed. Given a sequence $(\alpha_n)_{n=1}^{\infty}$ of real numbers, we form the sequence

$$\sup\{\alpha_k : k \ge n\}, \qquad n = 1, 2, 3, \ldots,$$

a nonincreasing sequence in $(-\infty, \infty]$. Being nonincreasing, the preceding sequence converges to a point in $[-\infty, \infty]$; its limit is called the limit superior of the original sequence and denoted by $\limsup_{n\to\infty} \alpha_n$. The notion of limit inferior is defined analogously and satisfies $\liminf_{n\to\infty} \alpha_n = -\limsup_{n\to\infty}(-\alpha_n)$. The sequence $(\alpha_n)_1^{\infty}$ converges finitely if and only if its limit superior and limit inferior are equal and finite.

Here are a few simple examples:

(1) $\limsup_{n\to\infty}\left(1+\frac{1}{n}\right) = \limsup_{n\to\infty}\left(1-\frac{1}{n}\right) = 1;$

(2) $\limsup_{n\to\infty}(-1)^n = 1;$

(3) $\limsup_{n\to\infty} n = \infty;$

(4) $\limsup_{n\to\infty}(-n) = -\infty.$

The following properties of limit superior can be found in any book on real analysis. Their proofs would be good exercises for the reader unfamiliar with the notion.

(i) If $(\alpha_n)_1^\infty$ and $(\beta_n)_1^\infty$ are two sequences of real numbers, then

$$\limsup_{n\to\infty}(\alpha_n+\beta_n) \le \limsup_{n\to\infty}\alpha_n + \limsup_{n\to\infty}\beta_n,$$

provided the sum on the right is well defined (i.e., excluding the case where one summand is ∞ and the other is $-\infty$). If one of the sequences converges then the equality holds (with the same proviso).

(ii) If $(\alpha_n)_1^\infty$ and $(\beta_n)_1^\infty$ are two sequences of positive real numbers, then

$$\limsup_{n\to\infty}(\alpha_n\beta_n) \le (\limsup_{n\to\infty}\alpha_n)(\limsup_{n\to\infty}\beta_n),$$

provided the product on the right is well defined (i.e., excluding the case where one factor is 0 and the other is ∞). If one of the sequences converges then the equality holds (with the same proviso).

(iii) If $(\alpha_n)_1^\infty$ is a sequence of real numbers, and the real number ρ satisfies $\rho < \limsup_{n\to\infty}\alpha_n$, then the inequality $\rho < \alpha_n$ holds for infinitely many indices n.

(iv) If $(\alpha_n)_1^\infty$ is a sequence of real numbers, and the real number ρ satisfies $\rho > \limsup_{n\to\infty}\alpha_n$, then the inequality $\rho > \alpha_n$ holds for all but perhaps finitely many indices n.

V.12. Cauchy-Hadamard Theorem

The radius of convergence of the power series $\sum_{n=0}^\infty a_n(z-z_0)^n$ *is equal to* $\dfrac{1}{\limsup_{n\to\infty}|a_n|^{\frac{1}{n}}}$.

The proof of this theorem is essentially the same as the proof in Section V.9. One could avoid the duplication of proofs by combining the Cauchy-Hadamard theorem and the proposition of V.9 into a single statement. The gain in efficiency, however, would be at the expense of digestibility.

It will obviously suffice to prove the theorem for the case $z_0 = 0$. Let R denote the radius of convergence of the power series $\sum_{n=0}^\infty a_n z^n$. We first establish the inequality $R \ge \dfrac{1}{\limsup_{n\to\infty}|a_n|^{\frac{1}{n}}}$, assuming that $\limsup_{n\to\infty}|a_n|^{\frac{1}{n}} < \infty$ (since otherwise the inequality is trivial). Let ρ be any positive number exceeding $\limsup_{n\to\infty}|a_n|^{\frac{1}{n}}$. By property (iv) in the preceding section, we then have $\rho > |a_n|^{\frac{1}{n}}$ except perhaps for finitely many indices n. Thus, with finitely many exceptions, the n-th coefficient of the series $\sum_{n=0}^{\infty} a_n z^n$ is dominated in absolute value by the n-th coefficient of the geometric series $\sum_{n=0}^{\infty} \rho^n z^n$, whose radius of convergence is $\frac{1}{\rho}$. It follows that $R \geq \frac{1}{\rho}$ (by Exercise V.10.1). Since ρ can be taken arbitrarily close to $\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$, we obtain the desired inequality, $R \geq \dfrac{1}{\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}}$.

To complete the proof we establish the opposite inequality,

$$R \leq \frac{1}{\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}},$$

assuming that $\limsup_{n\to\infty} |a_n|^{\frac{1}{n}} > 0$. Let ρ be any positive number less than $\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$. By property (iii) in the preceding section we then have $\rho < |a_n|^{\frac{1}{n}}$ for infinitely many indices n. Hence, if $|z| = \frac{1}{\rho}$, then $|a_n z^n| > 1$ for infinitely many indices n, implying that the series $\sum_{n=0}^{\infty} a_n z^n$ diverges. It follows that $R \leq \frac{1}{\rho}$, and since ρ can be taken arbitrarily close to $\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$, we obtain the desired inequality, $R \leq \dfrac{1}{\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}}$. The proof of the theorem is complete.

Exercise V.12.1. Let the power series $\sum_{n=0}^{\infty} a_n z^n$ and $\sum_{n=0}^{\infty} b_n z^n$ have respective radii of convergence R_1 and R_2. Prove that the radius of convergence of the series $\sum_{n=0}^{\infty} (a_n + b_n) z^n$ is at least the minimum of R_1 and R_2, and the radius of convergence of the series $\sum_{n=0}^{\infty} a_n b_n z^n$ is at least $R_1 R_2$ (provided the product is well defined).

Exercise* V.12.2. Prove that a power series $\sum_{n=0}^{\infty} a_n z^n$ and its termwise derivative, the series $\sum_{n=1}^{\infty} n a_n z^{n-1}$, have the same radius of convergence.

Exercise V.12.3. If the series $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R, what are the radii of convergence of the series $\sum_{n=0}^{\infty} a_n z^{2n}$ and $\sum_{n=0}^{\infty} a_n^2 z^n$?

V.13. Ratio Test

If $\lim_{n\to\infty} \left| \dfrac{a_n}{a_{n+1}} \right|$ *exists, then that limit is the radius of convergence of the power series* $\sum_{n=0}^{\infty} a_n (z - z_0)^n$. This will be deduced from the Cauchy-Hadamard theorem. Let R denote the radius of convergence of the power series, and let $R' = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|$ ($R' = \infty$ is allowed). We suppose first that R' is finite and prove that $R \le R'$. Let ρ be any positive number exceeding R'. Then the inequality $\left|\frac{a_n}{a_{n+1}}\right| < \rho$ holds for all except perhaps finitely many indices n, say for all $n \ge n_0$. We can assume $a_{n_0} \ne 0$ (since our hypothesis implies that only finitely many of the coefficients a_n can vanish). Iteration of the inequality $|a_{n+1}| > \rho^{-1}|a_n|$ gives

$$|a_{n_0+k}| > \rho^{-k}|a_{n_o}|, \qquad k = 1, 2, 3, \ldots.$$

Hence, for $n > n_0$ we have

$$|a_n|^{\frac{1}{n}} > \frac{1}{\rho}\rho^{\frac{n_0}{n}}|a_{n_0}|^{\frac{1}{n}}.$$

The right side in the preceding inequality tends to $\frac{1}{\rho}$ as $n \longrightarrow \infty$, implying (in view of the Cauchy-Hadamard theorem) that

$$\frac{1}{R} = \limsup_{n\to\infty}|a_n|^{\frac{1}{n}} \ge \frac{1}{\rho}.$$

Since ρ can be taken arbitrarily close to R', the desired inequality, $R \le R'$, follows.

To complete the proof we suppose $R' > 0$ and show that $R \ge R'$. The argument is nearly identical to the preceding one. Let ρ be any positive number less than R'. Then the inequality $\left|\frac{a_n}{a_{n+1}}\right| > \rho$ holds for all except perhaps finitely many indices n, say for all $n \ge n_0$. As above, we can iterate the inequality $|a_{n+1}| < \rho^{-1}|a_n|$ to get, for $n > n_0$,

$$|a_n|^{\frac{1}{n}} < \frac{1}{\rho}\rho^{\frac{n_0}{n}}|a_{n_0}|^{\frac{1}{n}}.$$

The quantity on the right side having the limit $\frac{1}{\rho}$ as $n \longrightarrow \infty$, we can conclude that

$$\frac{1}{R} = \limsup_{n\to\infty}|a_n|^{\frac{1}{n}} \le \frac{1}{\rho}.$$

Since ρ can be taken arbitrarily close to R', the desired inequality, $R \ge R'$, follows.

V.14. Examples

When it applies, the ratio test is generally easier to use than the Cauchy-Hadamard theorem. Often, however, the simple comparison test given in Exercise V.10.1 suffices to determine the radius of convergence.

Example 1. $\sum_{n=1}^{\infty} \frac{z^n}{n}$.

One can use the ratio test to show that the radius of convergence of the series is 1. Alternatively, a comparison with the series $\sum_{n=0}^{\infty} z^n$ shows that the radius of convergence is at least 1. The radius of convergence is at most 1 because the series diverges at $z = 1$.

Example 2. $\sum_{n=0}^{\infty} (n+1)z^n$.

Again, the ratio test gives immediately that the radius of convergence equals 1. The same conclusion can be reached from Exercise V.12.2, since the given series is the termwise derivative of the series $\sum_{n=0}^{\infty} z^n$.

Example 3. $\sum_{n=0}^{\infty} \frac{z^n}{n!}$.

Example 4. $\sum_{n=0}^{\infty} (n!)z^n$.

In each case the ratio test applies. In Example 3 the radius of convergence is ∞, in Example 4 it is 0.

Exercise V.14.1. Find the radius of convergence of each of the following series.

$$(a)\ \sum_{n=1}^{\infty} \frac{z^n}{n^3}; \qquad (b)\ \sum_{n=0}^{\infty} \frac{(n!)^3}{(3n)!} z^{3n}; \qquad (c)\ \sum_{n=1}^{\infty} \frac{z^{n!}}{n};$$

$$(d)\ \sum_{n=0}^{\infty} (n!)z^{n!}; \qquad (e)\ \sum_{n=1}^{\infty} n^n z^{n^2}.$$

Exercise V.14.2. Prove that the series $\sum_{n=1}^{\infty} \frac{z^n}{n}$ converges at every point of the unit circle except the point $z = 1$.

V.15. Differentiation of Power Series

Let the power series $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ have a positive radius of convergence R. Then the function that is represented in the disk $|z - z_0| < R$ by the series is holomorphic, and its derivative is represented in that disk by the series $\sum_{n=1}^{\infty} na_n(z - z_0)^{n-1}$, the termwise derivative of the original series.

An immediate corollary is that the function represented in the disk $|z - z_0| < R$ by the original series is differentiable to all orders, its k-th derivative being represented by the termwise k-th derivative of the original series. In particular, the value at z_0 of the k-th derivative of that function is $k!a_k$. A power series with center z_0 and positive radius of convergence R is thus the Taylor series with center z_0 of the function it represents.

In proving the proposition, we can assume without loss of generality that $z_0 = 0$. From Exercise V.12.2 we know that the series $\sum_{n=0}^{\infty} a_n z^n$ and $\sum_{n=1}^{\infty} n a_n z^{n-1}$ have the same radius of convergence. Let f and g denote the respective functions in the disk $|z| < R$ represented by these series. Fix a point z_1 in that disk. We shall prove that $f'(z_1) = g(z_1)$ by estimating the size of the difference between $\dfrac{f(z) - f(z_1)}{z - z_1}$ and $g(z_1)$ as z tends to z_1. Use will be made of the following standard algebraic identity, which holds for any two complex numbers z and w and any positive integer n:

$$(*) \qquad z^n - w^n = (z - w) \sum_{k=0}^{n-1} z^k w^{n-k-1}.$$

One proves this by multiplying out the right side and making the obvious cancellations. (The case $w = 1$ of $(*)$ appeared earlier in this chapter, in Section V.3, in the analysis of geometric series.)

Fix a positive number ρ such that $|z_1| < \rho < R$. Henceforth we assume that $|z| < \rho$. We have

$$\begin{aligned} \frac{f(z) - f(z_1)}{z - z_1} - g(z_1) &= \frac{1}{z - z_1}\left[\sum_{n=0}^{\infty} a_n z^n - \sum_{n=0}^{\infty} a_n z_1^n\right] - \sum_{n=1}^{\infty} n a_n z_1^{n-1} \\ &= \sum_{n=1}^{\infty} a_n \left[\frac{z^n - z_1^n}{z - z_1} - n z_1^{n-1}\right]. \end{aligned}$$

By $(*)$ the last expression can be rewritten as

$$\sum_{n=1}^{\infty} a_n \left[\left(\sum_{k=0}^{n-1} z^k z_1^{n-k-1}\right) - n z_1^{n-1}\right],$$

which is the same as

$$\sum_{n=2}^{\infty} a_n \left[\sum_{k=1}^{n-1} z_1^{n-k-1}\left(z^k - z_1^k\right)\right].$$

Using (*) again, we have

$$\begin{aligned}|z^k - z_1^k| &= |z - z_1|\left|\sum_{j=0}^{k-1} z^j z_1^{k-j-1}\right| \\ &\le |z - z_1|\sum_{j=0}^{k-1} |z|^j |z_1|^{k-j-1} \\ &\le k\rho^{k-1}|z - z_1|.\end{aligned}$$

Thus

$$\begin{aligned}\left|\frac{f(z)-f(z_1)}{z-z_1} - g(z_1)\right| &= \left|\sum_{n=2}^{\infty} a_n\left[\sum_{k=1}^{n-1} z_1^{n-k-1}\left(z^k - z_1^k\right)\right]\right| \\ &\le \sum_{n=2}^{\infty}|a_n|\left[\sum_{k=1}^{n-1}|z_1|^{n-k-1}|z^k - z_1^k|\right] \\ &\le \sum_{n=2}^{\infty}|a_n|\left[\sum_{k=1}^{n-1}\rho^{n-k-1}(k\rho^{k-1}|z-z_1|)\right] \\ &= |z - z_1|\sum_{n=2}^{\infty}|a_n|\left[\frac{n(n-1)}{2}\rho^{n-2}\right].\end{aligned}$$

The series $\sum_{n=2}^{\infty}\dfrac{n(n-1)}{2}|a_n|\rho^{n-2}$ converges, by V.9 in conjunction with Exercise V.12.2, applied twice. Letting M denote its sum, we can write our estimate as

$$\left|\frac{f(z)-f(z_1)}{z-z_1} - g(z_1)\right| \le M|z - z_1| \qquad (|z| < \rho).$$

We conclude that $\lim\limits_{z\to z_1}\dfrac{f(z)-f(z_1)}{z-z_1} = g(z_1)$; in other words, f is differentiable at z_1 and $f'(z_1) = g(z_1)$. The proof is complete.

V.16. Examples

So far, the only power series we have summed explicitly is the series $\sum_{n=0}^{\infty} z^n$. We are now able to sum a few more series.

Example 1. The series $\sum_{n=0}^{\infty}\frac{z^n}{n!}$ has radius of convergence ∞. It equals its termwise derivative. If f is the function the series represents then, by V.15, we have $f' = f$. By Exercise IV.3.1 we can conclude that f is a constant times e^z. But since the functions f and e^z both equal 1 at the origin, the constant must be 1. Thus $\sum_{n=0}^{\infty}\frac{z^n}{n!} = e^z$ for all z.

In a more sophisticated approach to complex analysis than is being attempted here, one uses the preceding equality to define e^z. By clever use of power series one then deduces all of the basic properties of the exponential function and of its relatives, the hyperbolic and trigonometric functions (without assuming prior knowledge of the real-variable versions of these functions). This approach can be found, for example, in the books of L. V. Ahlfors, and S. Saks and A. Zygmund, cited in the list of references.

Example 2. The power series $\sum_{n=1}^{\infty}\frac{z^n}{n}$ has radius of convergence 1. Its termwise derivative is the series $\sum_{n=0}^{\infty} z^n$. Hence, if f is the function in the disk $|z|<1$ represented by the series $\sum_{n=1}^{\infty}\frac{z^n}{n}$, then, by V.15, $f'(z)=\frac{1}{1-z}$. By Exercise II.8.1(a) we can conclude that the functions f and $\operatorname{Log}\frac{1}{1-z}$ differ by a constant. But both functions equal 0 at the origin, and hence they are equal:

$$\sum_{n=1}^{\infty}\frac{z^n}{n} = \operatorname{Log}\frac{1}{1-z} \qquad (|z|<1).$$

Exercise V.16.1. Find the power series with center z_0 that represents the function e^z.

Exercise V.16.2. What function is represented by the power series $\sum_{n=1}^{\infty} n^2 z^n$?

Exercise V.16.3. Let k be a nonnegative integer. Prove that the power series

$$\sum_{n=0}^{\infty}\frac{(-1)^n\left(\frac{z}{2}\right)^{k+2n}}{n!(n+k)!}$$

has radius of convergence ∞. The function represented by the series is called the Bessel function of the first kind of order k and denoted by J_k. Prove that J_k satisfies Bessel's differential equation:

$$z^2f''(z)+zf'(z)+(z^2-k^2)f(z)=0.$$

(Bessel functions, of which those introduced above are particular examples, have been exhaustively studied. They arise in numerous questions in mathematics and physics, for instance, in the study of a stretched circular membrane, to cite the most popular example.)

Exercise V.16.4. Find an expression for $\frac{\mathrm{d}}{\mathrm{d}z}\left[z^{-k}J_k(z)\right]$ in terms of $J_{k+1}(z)$.

V.17. Cauchy Product

Let $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ and $\sum_{n=0}^{\infty} b_n(z-z_0)^n$ be two power series with the same center. Their Cauchy product is by definition the power series $\sum_{n=0}^{\infty} c_n(z-z_0)^n$ whose n-th coefficient is given by $c_n=\sum_{k=0}^{n} a_k b_{n-k}$. It arises when one forms all products $a_j(z-z_0)^j\, b_k(z-z_0)^k$, adds for each n the ones with $j+k=n$, and sums the resulting terms. Suppose the series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ has a positive radius of convergence R_1 and the series $\sum_{n=0}^{\infty} b_n(z-z_0)^n$ has a positive radius of convergence R_2. We shall prove that their Cauchy product converges in the disk $|z - z_0| < \min\{R_1, R_2\}$ to the product of the functions represented by the two original series.

We can assume without loss of generality that $z_0 = 0$. Suppose $|z| < \min\{R_1, R_2\}$. For N a positive integer we have

$$\begin{aligned}
&\left(\sum_{j=0}^{N} a_j z^j\right)\left(\sum_{k=0}^{N} b_k z^k\right) - \sum_{n=0}^{N} c_n z^n \\
&= \sum_{0\le j,k\le N} a_j b_k z^{j+k} - \sum_{n=0}^{N}\sum_{j+k=n} a_j b_k z^{j+k} \\
&= \sum_{\substack{0\le j,k\le N\\ j+k>N}} a_j b_k z^{j+k}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\left|\left(\sum_{j=0}^{N} a_j z^j\right)\left(\sum_{k=0}^{N} b_k z^k\right) - \sum_{n=0}^{N} c_n z^n\right| \\
&\le \sum_{\substack{0\le j,k\le N\\ j+k>N}} |a_j b_k z^{j+k}| \\
&\le \sum_{\frac{N}{2}<\max\{j,k\}\le N} |a_j b_k z^{j+k}| \\
&\le \left(\sum_{j>\frac{N}{2}} |a_j z^j|\right)\left(\sum_{k=0}^{N} |b_k z^k|\right) + \left(\sum_{j=0}^{N} |a_j z^j|\right)\left(\sum_{k>\frac{N}{2}} |b_k z^k|\right) \\
&\le \left(\sum_{j>\frac{N}{2}} |a_j z^j|\right)\left(\sum_{k=0}^{\infty} |b_k z^k|\right) + \left(\sum_{j=0}^{\infty} |a_j z^j|\right)\left(\sum_{k>\frac{N}{2}} |b_k z^k|\right).
\end{aligned}$$

The last expression tends to 0 as $N \longrightarrow \infty$, because both series $\sum_{j=0}^{\infty} |a_j z^j|$ and $\sum_{k=0}^{\infty} |b_k z^k|$ converge. In view of the preceding inequality, therefore, we can conclude that

$$\sum_{n=0}^{\infty} c_n z^n = \left(\sum_{j=0}^{\infty} a_j z^j\right)\left(\sum_{k=0}^{\infty} b_k z^k\right),$$

as desired.

Exercise V.17.1. Deduce the law of exponents, $e^{z_1}e^{z_2} = e^{z_1+z_2}$, from the power series representation of e^z.

Exercise V.17.2. Consider the correspondence that associates with each power series $\sum_{n=0}^{\infty} a_n z^n$ centered at 0 the infinite matrix $(\alpha_{jk})_{j,k=0}^{\infty}$ defined by

$$\alpha_{jk} = \begin{cases} a_{j-k} & \text{for } j \geq k \\ 0 & \text{for } j < k. \end{cases}$$

The correspondence is obviously linear. Prove that it is also multiplicative, in other words, that the matrix associated with the Cauchy product of the two series is the product of the matrices associated with the two series. (The matrices arising in this exercise are instances of Toeplitz matrices, which is to say that they have constant diagonals.)

V.18. Division of Power Series

Suppose the power series $\sum_{n=0}^{\infty} b_n(z-z_0)^n$ and $\sum_{n=0}^{\infty} c_n(z-z_0)^n$ have positive radii of convergence and so represent holomorphic functions g and h, respectively, in disks with center z_0. Suppose also that $g(z_0) = b_0 \neq 0$. The quotient $f = h/g$ is then holomorphic in some disk with center z_0. Assume f is represented in that disk by a power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ (an assumption to be justified in VII.8). How does one find the coefficients a_n in terms of the coefficients b_n and c_n?

A method that always works in principle uses the Cauchy product, according to which

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}, \qquad n = 0, 1, \ldots.$$

From this we can conclude that $a_0 = c_0/b_0$ and

$$a_n = \frac{1}{b_0}\left(c_n - \sum_{k=0}^{n-1} a_k b_{n-k}\right), \qquad n = 1, 2, \ldots.$$

The last equality expresses a_n in terms of c_n, $b_0, \ldots, b_n$ and $a_0, \ldots, a_{n-1}$, enabling one to determine the coefficients a_n recursively starting from the initial value $a_0 = c_0/b_0$.

Exercise V.18.1. Use the scheme above to determine the power series with center 0 representing the function $f(z) = 1/(1+z+z^2)$ near 0. What is the radius of convergence of the series?

Complex Integration

The integral of most use in complex function theory is a type of line integral, usually taken around a closed curve in **C**. The complex line integral is introduced, and its basic properties developed, in the present chapter. The first three sections contain preliminary material on integration of complex-valued functions of a real variable.

VI.1. Riemann Integral for Complex-Valued Functions

The complex-valued function ϕ on the closed subinterval $[a, b]$ of **R** is said to be piecewise continuous if it is continuous at all but finitely many points of $[a, b]$ and has finite one-sided limits at each point of discontinuity. In that case the Riemann integrals of Re ϕ and Im ϕ over $[a, b]$ are defined, and we define the integral of ϕ over $[a, b]$ by

$$\int_a^b \phi(t)dt = \int_a^b \operatorname{Re}\phi(t)dt + i\int_a^b \operatorname{Im}\phi(t)dt.$$

From the linearity of the Riemann integral for real-valued functions one obtains immediately the linearity for complex-valued functions: if ϕ_1 and ϕ_2 are piecewise-continuous complex-valued functions on $[a, b]$ and c_1 and c_2 are complex constants, then

$$\int_a^b \big[c_1\phi_1(t) + c_2\phi_2(t)\big]dt \;=\; c_1\int_a^b \phi_1(t)dt + c_2\int_a^b \phi_2(t)dt.$$

Of course, one can use the Riemann integral under less restrictive conditions than the piecewise continuity of the function to be integrated, but we shall not need to do so here.

VI.2. Fundamental Theorem of Calculus

The complex-valued function ϕ on the interval $[a,b]$ is said to be differentiable at the point t_0 of $[a,b]$ if Re ϕ and Im ϕ are both differentiable at t_0; we then define the derivative of ϕ at t_0 by $\phi'(t_0) = (\operatorname{Re}\phi)'(t_0) + i(\operatorname{Im}\phi)'(t_0)$. (This notion was used earlier, in Section II.10.) The function ϕ is said to be piecewise C^1 if it is continuous, it is differentiable at all but finitely many points of $[a,b]$, its derivative is continuous at each point where it exists, and its derivative has finite one-sided limits at each of the finitely many points where it fails to exist. From the fundamental theorem of calculus for real-valued functions one obtains immediately the corresponding result for complex-valued functions: if the complex-valued function ϕ on $[a,b]$ is piecewise C^1, then

$$\int_a^b \phi'(t)dt \;=\; \phi(b) - \phi(a).$$

Exercise VI.2.1. Prove that a piecewise-C^1 function has one-sided derivatives at each point where it is not differentiable.

VI.3. Triangle Inequality for Integration

If ϕ is a piecewise-continuous complex-valued function on the interval $[a,b]$, then

$$\left|\int_a^b \phi(t)dt\right| \;\le\; \int_a^b |\phi(t)|dt.$$

To establish this, assume $\int_a^b \phi(t)dt \neq 0$ (otherwise the inequality is trivial), and let

$$\lambda \;=\; \frac{\left|\int_a^b \phi(t)dt\right|}{\int_a^b \phi(t)dt}.$$

Then

$$\begin{aligned}
\left|\int_a^b \phi(t)dt\right| &= \lambda \int_a^b \phi(t)dt \\
&= \int_a^b \lambda\phi(t)dt \;=\; \operatorname{Re}\int_a^b \lambda\phi(t)dt \\
&= \int_a^b \operatorname{Re}\left[\lambda\phi(t)\right]dt \;\le\; \int_a^b |\lambda\phi(t)|dt \\
&= \int_a^b |\phi(t)|dt,
\end{aligned}$$

as desired.

VI.4. Arc Length

Although the reader has no doubt already encountered the notion of arc length, a brief review is in order. The length of the piecewise-C^1 curve $\gamma : [a,b] \longrightarrow \mathbf{C}$ is the number

$$L(\gamma) = \int_a^b |\gamma'(t)|dt.$$

The reader who wishes may take the preceding equality as the definition of $L(\gamma)$. It is more natural, however, to define $L(\gamma)$, for any curve γ (not necessarily piecewise C^1) to be the supremum of the lengths of all inscribed polygonal paths. The integral expression above for $L(\gamma)$, in the case where γ is piecewise C^1, then becomes a theorem. Details can be found in the book of W. Rudin cited in the list of references.

One can gain an intuitive understanding of arc length by interpreting the parameter t as time. One then thinks of γ as describing the position of a point moving in the plane. In case γ is piecewise C^1, the derivative $\gamma'(t)$, at a value t where it exists, gives the instantaneous velocity of the moving point, and $|\gamma'(t)|$ gives the instantaneous speed. The total distance the point covers between the instants $t = a$ and $t = b$ should thus be the integral of $|\gamma'|$ over $[a,b]$.

VI.5. The Complex Integral

Let $\gamma : [a,b] \longrightarrow \mathbf{C}$ be a piecewise-C^1 curve. A complex-valued function f is said to be defined on γ if it is defined on a set that contains the range of γ. In that case, the composite function $f \circ \gamma$ is defined. If f is continuous and defined on γ, then $f \circ \gamma$ is continuous, and we define the integral of f over γ, denoted $\int_\gamma f(z)dz$, by

$$\int_\gamma f(z)dz = \int_a^b f(\gamma(t))\gamma'(t)dt.$$

The integral just defined has the usual linearity property: if f_1 and f_2 are two continuous functions defined on γ, and if c_1 and c_2 are complex constants, then

$$\int_\gamma \left[c_1 f_1(z) + c_2 f_2(z)\right]dz = c_1 \int_\gamma f_1(z)dz + c_2 \int_\gamma f_2(z)dz.$$

This follows immediately from the analogous property mentioned in Section VI.1.

Another simple property is that we can partition γ into two (or more) subcurves and then express $\int_\gamma f(z)dz$ as the sum of integrals over the subcurves. For example, if $a < c < b$, and if $\gamma_1 = \gamma\big|[a,c]$ and $\gamma_2 = \gamma\big|[c,b]$, then $\int_\gamma f(z)dz = \int_{\gamma_1} f(z)dz + \int_{\gamma_2} f(z)dz$. This is a trivial consequence of the corresponding property of the Riemann integral.

Sometimes we shall wish to use a symbol other than "z" to denote the variable of integration (the "dummy" variable). A frequent choice is "ζ". There is no difference between $\int_\gamma f(\zeta)d\zeta$ and $\int_\gamma f(z)dz$. Both by definition equal $\int_a^b f(\gamma(t))\gamma'(t)dt$. (This is the same sort of flexibility in notation one encounters when first studying integration, in calculus.)

VI.6. Integral of a Derivative

Let $\gamma : [a, b] \longrightarrow \mathbf{C}$ be a piecewise-C^1 curve. Let f be a holomorphic function defined in an open set containing γ, and assume that f' is continuous. Then

$$\int_\gamma f'(z)dz = f(\gamma(b)) - f(\gamma(a)).$$

In particular, if γ is closed (i.e., $\gamma(b) = \gamma(a)$), then $\int_\gamma f'(z)dz = 0$.

In fact, by definition we have

$$\int_\gamma f'(z)dz = \int_a^b f'(\gamma(t))\,\gamma'(t)dt.$$

By the chain rule (see Appendix 2), the integrand in the integral on the right side is the derivative of the composite function $f \circ \gamma$. The desired conclusion thus follows from the version of the fundamental theorem of calculus given in Section VI.2.

As mentioned in Chapter II, we shall prove later that a holomorphic function has a holomorphic derivative, making superfluous the assumption in the preceding proposition that f' is continuous. For the time being, however, that assumption is needed. If f is represented by a power series, convergent in the whole plane, say, then we can use the proposition, for we know from Chapter V that f is then holomorphic and that f' is represented by a power series, and hence is continuous.

VI.7. An Example

Let z_0 be a point of $\mathbf{C}$ and R a positive real number. Define $\gamma : [0, 2\pi] \longrightarrow \mathbf{C}$ by $\gamma(t) = z_0 + Re^{it}$. Thus, γ is a parametrization of the circle $|z - z_0| = R$, traversed once in the counterclockwise direction. If $n \neq -1$ then, by VI.6, $\int_\gamma (z - z_0)^n dz = 0$, because $(z - z_0)^n$ is the derivative of $\frac{(z-z_0)^{n+1}}{n+1}$.

On the other hand, referring back to the definition of the integral, we have

$$\int_\gamma (z - z_0)^{-1}dz = \int_0^{2\pi} (Re^{it})^{-1} Rie^{it}dt = i\int_0^{2\pi} dt = 2\pi i.$$

Exercise VI.7.1. Prove on the basis of the last calculation that there is no branch of $\arg z$ in the region $0 < |z| < 1$.

Exercise VI.7.2. Derive the formula

$$\frac{1}{2\pi}\int_0^{2\pi} \cos^{2n} t\,dt \;=\; \frac{1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1)}{2\cdot 4\cdot 6\cdot\cdots\cdot(2n)}$$

by integrating the function $\frac{1}{z}\left(z+\frac{1}{z}\right)^{2n}$ around the unit circle, parametrized by the curve $\gamma(t) = e^{it}$ $(0 \le t \le 2\pi)$. (Suggestion: Expand the binomial.)

VI.8. Reparametrization

Let $\gamma : [a,b] \longrightarrow \mathbf{C}$ be a piecewise-C^1 curve. One says that the curve $\gamma_1 : [a_1, b_1] \longrightarrow \mathbf{C}$ is a reparametrization of γ if it has the form $\gamma_1 = \gamma\circ\beta$, where β is an increasing piecewise-C^1 map of $[a_1, b_1]$ onto $[a,b]$. In this case, if f is a continuous function defined on γ (and hence on γ_1), then

$$\int_\gamma f(z)dz \;=\; \int_{\gamma_1} f(z)dz.$$

In fact, from the definition of the complex integral,

$$\int_\gamma f(z)dz \;=\; \int_a^b f(\gamma(t))\gamma'(t)dt,$$

$$\int_{\gamma_1} f(z)dz \;=\; \int_{a_1}^{b_1} f(\gamma(\beta(s)))\gamma'(\beta(s))\beta'(s)ds.$$

The two integrals on the right side are equal by the change-of-variables formula for the Riemann integral.

Thus, the curve γ and its reparametrizations are indistinguishable for purposes of integration. This will often save us the trouble, when we wish to perform an integration, of mentioning a particular parametrization; it will usually be enough to give a geometric description of the integration curve and to specify the direction of integration.

For example, if z_1 and z_2 are distinct points of $\mathbf{C}$, then by $[z_1, z_2]$ we shall mean the segment with endpoints z_1 and z_2, directed from z_1 to z_2. The standard parametrization of the segment is given by the function $\gamma(t) = (1-t)z_1 + tz_2$ $(0 \le t \le 1)$. If the function f is defined and continuous on $[z_1, z_2]$, then $\int_{[z_1,z_2]} f(z)dz$ has an unambiguous meaning: it equals $\int_\gamma f(z)dz$ for the γ above or any of its reparametrizations. We need not be explicit about a parametrization to be able to talk about $\int_{[z_1,z_2]} f(z)dz$.

For closed curves we have even a bit more flexibility regarding parametrizations, in that the choice of initial-terminal point is irrelevant for purposes of integration. Suppose $\gamma : [a,b] \longrightarrow \mathbf{C}$ is a piecewise-C^1 curve with

$\gamma(a) = \gamma(b)$, and let a_1 be a point in (a, b). Let $b_1 = a_1 + (b - a)$, and define the curve $\gamma_1 : [a_1, b_1] \longrightarrow \mathbf{C}$ by

$$\gamma_1(t) = \begin{cases} \gamma(t), & a_1 \leq t \leq b \\ \gamma(t + a - b), & b \leq t \leq b_1. \end{cases}$$

Thus, we obtain γ_1 by "shifting" γ so as to make $\gamma(a_1)$, rather than $\gamma(a)$, the initial-terminal point of the parametrization. The reader will easily verify that $\int_\gamma f(z)dz = \int_{\gamma_1} f(z)dz$ for any function f that is defined and continuous on γ.

Exercise VI.8.1. Let z_1 and z_2 be distinct points of $\mathbf{C}$. Evaluate $\int_{[z_1,z_2]} z^n dz$ and $\int_{[z_1,z_2]} \bar{z}^n dz$ for $n = 0, 1, 2, \ldots$.

Exercise VI.8.2.

(a) Evaluate $\int_{[-i,i]} |z| dz$.

(b) Evaluate $\int_{C_+} |z| dz$, where C_+ is the right half of the unit circle, oriented counterclockwise.

Exercise VI.8.3. Let the complex-valued function f be defined and continuous in the disk $|z - z_0| < R$. For $0 < r < R$ let C_r denote the circle $|z - z_0| = r$, with the counterclockwise orientation. Prove that

$$\lim_{r \to 0} \int_{C_r} \frac{f(z)}{z - z_0} dz = 2\pi i\, f(z_0).$$

Exercise VI.8.4. Let f, R and C_r be as in the preceding exercise, and assume that f is of class C^1. Prove that

$$\lim_{r \to 0} \frac{1}{r^2} \int_{C_r} f(z)dz = 2\pi i \frac{\partial f}{\partial \bar{z}}(z_0).$$

(See Section II.16 for the definition of $\frac{\partial f}{\partial \bar{z}}$.) Conclude that f is differentiable at z_0 if and only if

$$\lim_{r \to 0} \frac{1}{r^2} \int_{C_r} f(z)dz = 0.$$

VI.9. The Reverse of a Curve

If $\gamma : [a, b] \longrightarrow \mathbf{C}$ is a curve, then the reverse of γ is the curve $-\gamma : [-b, -a] \longrightarrow \mathbf{C}$ defined by $(-\gamma)(t) = \gamma(-t)$. It is the curve one obtains from γ by reversing its direction. If γ is piecewise C^1, then so is $-\gamma$, and $\int_{-\gamma} f(z)dz = -\int_\gamma f(z)dz$ for any function f that is defined and continuous on γ. The reader will easily verify this on the basis of the change- of-variables formula for the Riemann integral.

VI.10. Estimate of the Integral

Let $\gamma : [a, b] \longrightarrow \mathbf{C}$ be a piecewise-C^1 curve and f a continuous complex-valued function on γ. Let M be the maximum of $|f|$ on γ, that is,

$$M = \max\{|f(\gamma(t))| : a \le t \le b\}.$$

Then

$$\left|\int_\gamma f(z)dz\right| \le M\,L(\gamma).$$

In fact, by VI.3 we have

$$\begin{aligned}\left|\int_\gamma f(z)dz\right| &= \left|\int_a^b f(\gamma(t))\gamma'(t)dt\right| \\ &\le \int_a^b |f(\gamma(t))|\,|\gamma'(t)|dt.\end{aligned}$$

The last integral is majorized by $M\int_a^b |\gamma'(t)|dt$, which equals $M\,L(\gamma)$.

VI.11. Integral of a Limit

Let $\gamma : [a, b] \longrightarrow \mathbf{C}$ be a piecewise C^1 curve. Let the sequence $(f_n)_{n=1}^\infty$ of continuous complex-valued functions on γ converge uniformly on γ to the function f. Then

$$\lim_{n\to\infty}\int_\gamma f_n(z)dz = \int_\gamma f(z)dz.$$

In fact, let M_n denote the maximum of $|f - f_n|$ on γ. By VI.10,

$$\left|\int_\gamma f(z)dz - \int_\gamma f_n(z)dz\right| = \left|\int_\gamma [f(z) - f_n(z)]dz\right| \le M_n L(\gamma).$$

Since $\lim_{n\to\infty} M_n = 0$ (by the hypothesis that $f_n \longrightarrow f$ uniformly on γ), the desired conclusion follows.

VI.12. An Example

It is now possible to illustrate how complex methods can be used to evaluate improper Riemann integrals that are inaccessible by the usual techniques of calculus. Many other examples of this kind will be given later, in Chapters VII and X, after more machinery has been developed.

For a and b positive numbers, let $R_{a,b}$ denote the rectangle with vertices $\pm a$, $\pm a + ib$, oriented counterclockwise. By VI.6,

$$\int_{R_{a,b}} e^{-z^2}dz = 0.$$

We are going to hold b fixed and let a tend to ∞. The preceding integral is the sum of the integrals of e^{-z^2} over the four segments $[-a,a]$, $[a,a+ib]$, $[a+ib,-a+ib]$, $[-a+ib,-a]$; let those integrals be denoted by $I_1(a)$, $I_2(a)$, $I_3(a)$, $I_4(a)$, respectively. Obviously,

$$I_1(a) = \int_{-a}^{a} e^{-x^2}\,dx.$$

Parametrizing the segment $[-a+ib, a+ib]$ by the function $\gamma(t) = t+ib$, where $-a \le t \le a$, we have

$$\begin{aligned} I_3(a) &= \int_{-\gamma} e^{-z^2}dz = -\int_{\gamma} e^{-z^2}dz \\ &= -\int_{-a}^{a} e^{-(t+ib)^2}dt \\ &= -\int_{-a}^{a} e^{b^2-t^2}(\cos 2bt - i\sin 2bt)dt \\ &= -e^{b^2}\int_{-a}^{a} e^{-t^2}\cos 2bt\,dt. \end{aligned}$$

(The imaginary part disappeared because $e^{-t^2}\sin 2bt$ is an odd function of t.) Since $\sum_{j=1}^{4} I_j(a) = 0$, we obtain

$$e^{b^2}\int_{-a}^{a} e^{-t^2}\cos 2bt\,dt = \int_{-a}^{a} e^{-x^2}dx + I_2(a) + I_4(a).$$

Because $|e^{-z^2}| = e^{y^2-x^2}$, the maximum of $|e^{-z^2}|$ on the segment $[a, a+ib]$ is $e^{-a^2+b^2}$. By VI.10, therefore, $|I_2(a)| \le be^{-a^2+b^2}$. Exactly the same estimate holds for $I_4(a)$. Therefore $I_2(a)$ and $I_4(a)$ both tend to 0 as a tends to ∞. Taking the limit as $a \longrightarrow \infty$ in the equality above, we thus obtain

$$e^{b^2}\int_{-\infty}^{\infty} e^{-t^2}\cos 2bt\,dt = \int_{-\infty}^{\infty} e^{-x^2}dx.$$

Now, as the reader probably knows, $\int_{-\infty}^{\infty} e^{-x^2}dx$ can be found by elementary means. The trick is to note that

$$\left(\int_{-\infty}^{\infty} e^{-x^2}dx\right)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2}dx\,dy.$$

Switching to polar coordinates we can rewrite the preceding double integral as

$$\int_{0}^{\infty}\int_{0}^{2\pi} e^{-r^2} r\,d\theta\,dr,$$

which is easily seen to have the value π. Hence $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$. Combining this with the equality at the end of the last paragraph, we see that we have succeeded in evaluating the nonelementary integral $\int_{-\infty}^{\infty} e^{-t^2} \cos 2bt\, dt$:

$$\int_{-\infty}^{\infty} e^{-t^2} \cos 2bt\, dt = \sqrt{\pi} e^{-b^2}.$$

(The integral arises in Fourier analysis.)

Exercise VI.12.1. Prove that

$$\int_{0}^{\infty} e^{-t^2} \cos t^2 dt = \frac{1}{4}\sqrt{\pi}\sqrt{1+\sqrt{2}}$$

by integrating the function e^{-z^2} in the counterclockwise direction around the boundary of the region $\{z : |z| \le R,\ 0 \le \operatorname{Arg} z \le \frac{\pi}{8}\}$, and letting $R \longrightarrow \infty$.

Exercise VI.12.2. Evaluate the integrals $\int_0^\infty \cos t^2 dt$ and $\int_0^\infty \sin t^2 dt$ (the Fresnel integrals) by integrating e^{-z^2} in the counterclockwise direction around the boundary of the region $\{z : |z| \le R,\ 0 \le \operatorname{Arg} z \le \frac{\pi}{4}\}$, and letting $R \longrightarrow \infty$.

Chapter VII

Core Versions of Cauchy's Theorem, and Consequences

Cauchy's theorem is the central theorem of complex function theory. It states, in one of its versions, that $\int_\gamma f(z)dz = 0$ whenever γ is a closed curve and the function f is holomorphic in an open set containing γ and its interior. In a more sophisticated version, the single curve γ is replaced by a system of curves.

A general version of Cauchy's theorem is presented in Chapter IX; it requires considerable preparation. In the present chapter two special cases will be established, the case where γ is a triangle, and, on the basis of that, the case where the function f is holomorphic in a convex open set containing γ. This will be enough to enable us to draw far-reaching conclusions concerning the behavior of holomorphic functions. Some applications to harmonic functions are given at the end of the chapter.

VII.1. Cauchy's Theorem for a Triangle

Let T be a triangle in $\mathbf{C}$, and let f be a holomorphic function in an open set containing T and its interior. Then $\int_T f(z)dz = 0$.

This is sometimes called Goursat's lemma, after E. Goursat, who, in the late 1800's, was the first to recognize that, in the definition of a holomorphic function, one can dispense with the requirement that the derivative be continuous.

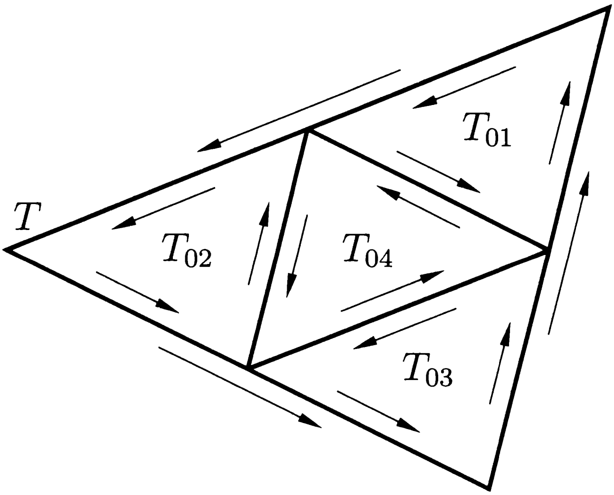

Figure 4. Carving up the triangle T

To prove the theorem, we assume for definiteness that T has the counterclockwise orientation, and we let $I = \int_T f(z)dz$. By means of segments joining the midpoints of the sides of T, we construct four triangles T_{0j}, $j = 1, 2, 3, 4$, each similar to but one-half the size of T (see Figure 4). Orienting each T_{0j} counterclockwise, we have

$$\int_T f(z)dz = \sum_{j=1}^{4} \int_{T_{0j}} f(z)dz.$$

Hence there is a j such that

$$\left| \int_{T_{0j}} f(z)dz \right| \geq \frac{1}{4}|I|.$$

Pick such a j, and let T_1 denote T_{0j} for that j.

From T_1 we construct four triangles T_{1j}, $j = 1, 2, 3, 4$, in the same way that the triangles T_{0j} were constructed from T. Orienting each T_{1j} counterclockwise, we have

$$\int_{T_1} f(z)dz = \sum_{j=1}^{4} \int_{T_{1j}} f(z)dz,$$

so, since $\left|\int_{T_1} f(z)dz\right| \geq \frac{1}{4}\left|I\right|$, there must be a j such that

$$\left|\int_{T_{1j}} f(z)dz\right| \geq \frac{1}{4^2}|I|.$$

Pick such a j, and let T_2 denote T_{1j} for that j.

Proceeding in this manner, we obtain a sequence $\left(T_n\right)_{n=1}^{\infty}$ of triangles, each T_n being one of the four triangles obtained from T_{n-1} by the process of connecting midpoints, such that

$$\left|\int_{T_n} f(z)dz\right| \geq \frac{1}{4^n}|I|$$

for each n. We note that $L\left(T_n\right) = \frac{1}{2^n}L\left(T\right)$ for each n.

Let K_n be the union of T_n and its interior. The sets K_n are compact, they decrease as n increases, and their diameters tend to 0. Hence, their intersection consists of a single point, which we call z_0. The point z_0 is either on T or in the interior of T, and thus it is in the open set where f is holomorphic. Hence, by II.2, we can write

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + R(z),$$

where $\lim_{z \to z_0} \dfrac{R(z)}{z - z_0} = 0$. By VI.6 the integral of the function $f(z_0)+f'(z_0)(z-z_0)$ around T_n is 0, so

$$\int_{T_n} f(z)dz = \int_{T_n} R(z)dz.$$

We thus have

$$|I| \leq 4^n\left|\int_{T_n} R(z)dz\right|.$$

For each n, let ϵ_n denote the maximum of $\left|\dfrac{R(z)}{z - z_0}\right|$ for z in $T_n\backslash\{z_0\}$. Then $\lim_{n\to\infty} \epsilon_n = 0$, and for z in T_n,

$$\left|R(z)\right| \leq \epsilon_n \operatorname{diameter}\left(T_n\right) \leq \epsilon_n\, L\left(T_n\right).$$

Hence

$$\begin{aligned}\left|\int_{T_n} R(z)dz\right| &\leq \epsilon_n\, L\left(T_n\right) \cdot L\left(T_n\right)\\ &= \frac{\epsilon_n}{4^n}L(T)^2.\end{aligned}$$

In combination with our earlier inequality, this gives

$$\left|I\right| \leq \epsilon_n\, L(T)^2.$$

Since $\lim_{n\to\infty} \epsilon_n = 0$, it follows that $I = 0$, as desired.

VII.2. Cauchy's Theorem for a Convex Region

Let G be a convex open subset of $\mathbf{C}$ *and f a holomorphic function in G. Then $\int_\gamma f(z)dz = 0$ for every piecewise-C^1 closed curve γ in G.*

We recall that a subset of $\mathbf{C}$ is called convex if, whenever it contains two points z_1 and z_2, it contains also the entire segment with endpoints z_1 and z_2. To prove the theorem we fix a point z_0 in G and define the function g in G by

$$g(z) = \int_{[z_0,z]} f(\zeta)d\zeta.$$

We shall show that g is holomorphic and that $g' = f$. The desired conclusion will then follow by VI.6.

Fix a point z_1 in G. If z is any other point of G, then

$$\int_{[z_0,z_1]} f(\zeta)d\zeta + \int_{[z_1,z]} f(\zeta)d\zeta + \int_{[z,z_0]} f(\zeta)d\zeta = 0,$$

by Cauchy's theorem for a triangle (or trivially in case the points z_0, z_1, z happen to be collinear). The preceding equality can be written as

$$g(z) - g(z_1) = \int_{[z_1,z]} f(\zeta)d\zeta.$$

Consequently

$$\frac{g(z) - g(z_1)}{z - z_1} - f(z_1) = \frac{1}{z - z_1}\int_{[z_1,z]}\Big[f(\zeta) - f(z_1)\Big]d\zeta.$$

Now fix a positive number ϵ. By the continuity of f, there is a positive number δ such that $\Big|f(\zeta) - f(z_1)\Big| < \epsilon$ whenever $\Big|\zeta - z_1\Big| < \delta$. Hence, if $\Big|z - z_1\Big| < \delta$, then the absolute value of the integrand in the preceding integral is bounded by ϵ, and we have

$$\begin{aligned}\left|\frac{g(z) - g(z_1)}{z - z_1} - f(z_1)\right| &= \frac{1}{\Big|z - z_1\Big|}\left|\int_{[z_1,z]}\Big[f(\zeta) - f(z_1)\Big]d\zeta\right| \\ &\leq \frac{1}{\Big|z - z_1\Big|}\cdot\epsilon L\Big(\big[z_1, z\big]\Big) = \epsilon.\end{aligned}$$

This proves that $\displaystyle\lim_{z\to z_1}\frac{g(z) - g(z_1)}{z - z_1} = f(z_1)$, in other words, g is differentiable at z_1 and $g'(z_1) = f(z_1)$, as desired.

VII.3. Existence of a Primitive

If the function f is defined in the open subset G of $\mathbf{C}$, then by a primitive of f one means a holomorphic function in G whose derivative equals f. The preceding proof shows that f has a primitive in case it is holomorphic and G is convex. The proof in fact establishes something that seems stronger: *If the continuous complex-valued function f in the open convex set G has integral 0 around every triangle contained in G, then f has a primitive in G.* This is not in reality stronger, because, as we shall show very shortly, the derivative of a holomorphic function is itself holomorphic, which means a function f satisfying the hypotheses in the preceding statement is in fact holomorphic. The preceding statement thus amounts to a converse of Goursat's lemma. The converse is stated formally in Section VII.10.

VII.4. More Definite Integrals

Cauchy's theorem for a convex region, although very special compared to the general Cauchy theorem, is by itself enough for many applications. In the following exercises it is used to evaluate improper integrals.

Exercise VII.4.1. Let $0 < b < 1$. Derive the equality

$$\int_{-\infty}^{\infty} \frac{1-b+x^2}{(1-b+x^2)^2+4bx^2}dx = \pi$$

by integrating the function $(1+z^2)^{-1}$ around the rectangle with vertices $\pm a$, $\pm a + i\sqrt{b}$ $(a > 0)$ and taking the limit as $a \to \infty$. What goes wrong if $b > 1$?

Exercise VII.4.2. Let a be a positive number. Evaluate the integrals

$$\int_0^{\infty} \frac{1}{t^4+a^4}dt, \qquad \int_0^{\infty} \frac{t^2}{t^4+a^4}dt$$

by integrating the function $(z^2+a^2)^{-1}$ in the counterclockwise direction around the boundary of the region $\{z : 0 \le |z| \le R,\ 0 \le \operatorname{Arg} z \le \frac{\pi}{4}\}$, and taking the limit as $R \to \infty$.

VII.5. Cauchy's Formula for a Circle

Let C be a counterclockwise oriented circle, and let f be a holomorphic function defined in an open set containing C and its interior. Then

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - z}d\zeta$$

for all points z in the interior of C.

Thus, the values of f in the interior of C can be reconstructed from the values on C. A holomorphic function "hangs together".

Cauchy's formula for a circle is a very special case of a more general formula, to be established in Chapter X. The case of a circle is all we shall

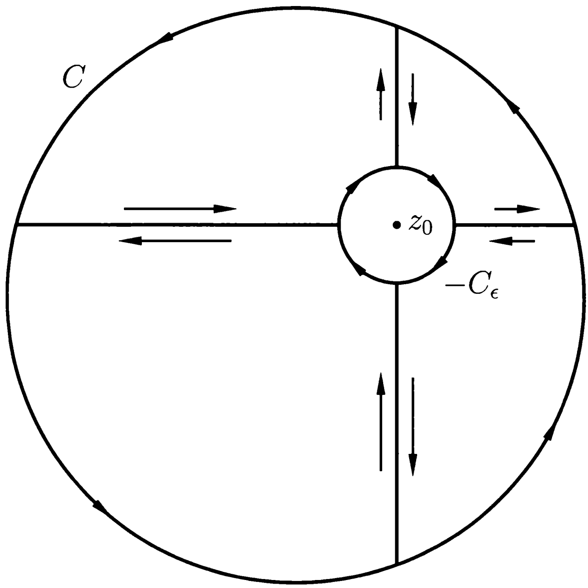

Figure 5. Integration curves in the proof of Cauchy's formula for a circle.

need to show that holomorphic functions can be represented locally by power series and thus are differentiable to all orders.

To establish the formula we fix a point z_0 in the interior of C. For $0 < \epsilon < \operatorname{dist}(z_0, C)$, let C_ϵ be the circle with center z_0 and radius ϵ, oriented counterclockwise. Our first task is to prove that $\int_C \frac{f(\zeta)}{\zeta - z_0} d\zeta = \int_{C_\epsilon} \frac{f(\zeta)}{\zeta - z_0} d\zeta$ for every ϵ.

The preceding equality will be deduced from Cauchy's theorem for a convex region. For each ϵ we subdivide the region between C and C_ϵ into four subregions by means of two horizontal and two vertical segments connecting C with C_ϵ, as in Figure 5. The boundary of each subregion is a closed curve made up of a subarc of C, a subarc of C_ϵ, and two of the segments, one horizontal and one vertical. Let these boundary curves be denoted by γ_1, γ_2, γ_3, γ_4, and let them be given the counterclockwise orientation. Each γ_j is contained in a convex open set in which the function $\frac{f(\zeta)}{\zeta - z_0}$ is holomorphic (for example, in the intersection of a suitable open half-plane with an open disk containing C in which f is holomorphic). By Cauchy's theorem for a convex region, therefore,

$$\int_{\gamma_j} \frac{f(\zeta)}{\zeta - z_0} d\zeta = 0, \quad j = 1, 2, 3, 4.$$

Since

$$\sum_{j=1}^{4}\int_{\gamma_j}\frac{f(\zeta)}{\zeta - z_0}d\zeta \;=\; \int_C \frac{f(\zeta)}{\zeta - z_0}d\zeta - \int_{C_\epsilon}\frac{f(\zeta)}{\zeta - z_0}d\zeta,$$

we have the desired equality, $\displaystyle\int_C \frac{f(\zeta)}{\zeta - z_0}d\zeta = \int_{C_\epsilon}\frac{f(\zeta)}{\zeta - z_0}d\zeta.$

In Section VI.7 it is shown that

$$\int_{C_\epsilon}\frac{1}{\zeta - z_0}d\zeta \;=\; 2\pi i.$$

This in conjunction with the preceding equality gives

$$\frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - z_0}d\zeta \;-\; f(z_0) \;=\; \frac{1}{2\pi i}\int_{C_\epsilon}\frac{f(\zeta) - f(z_0)}{\zeta - z_0}d\zeta$$

for each ϵ between 0 and $\operatorname{dist}(z_0, C)$. Let M_ϵ denote the maximum of $\left|\dfrac{f(\zeta) - f(z_0)}{\zeta - z_0}\right|$ for ζ on C_ϵ. The absolute value of the last integral is then bounded by $M_\epsilon\, L(C_\epsilon)$, implying that

$$\left|\frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - z_0}d\zeta \;-\; f(z_0)\right| \;\le\; \epsilon M_\epsilon.$$

As ϵ tends to 0 so does $\epsilon\, M_\epsilon$ (since $M_\epsilon \longrightarrow |f'(z_0)|$), and the desired equality follows.

Exercise VII.5.1. Let f and C be as in the statement of Cauchy's formula. What is the value of $\displaystyle\int_C \frac{f(\zeta)}{\zeta - z}d\zeta$ when z is in the exterior of C?

VII.6. Mean Value Property

Let the function f be holomorphic in the disk $|z - z_0| < R$, and for $0 < r < R$ let C_r denote the circle with center z_0 and radius r, oriented counterclockwise. By Cauchy's formula for a circle we have

$$f(z_0) \;=\; \frac{1}{2\pi i}\int_{C_r}\frac{f(z)}{z - z_0}dz, \quad 0 < r < R.$$

Parametrizing C_r by means of the function $z_0 + re^{it}$ $(0 \le t \le 2\pi)$, and rewriting the complex integral as an integral with respect to the parameter t, we obtain the equality

$$f(z_0) \;=\; \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{it})dt.$$

In other words, the value of f at the center of C_r is the average of its values over C_r.

Exercise VII.6.1. Retaining the notations above, and writing $z = x + iy$, prove that

$$f(z_0) = \frac{1}{\pi r^2} \int\int_{|z-z_0|<r} f(x+iy)dxdy, \quad 0 < r < R;$$

in other words, the value of f at the center of the disk $|z - z_0| < r$ is the average of its values over the disk.

VII.7. Cauchy Integrals

Let $\gamma : [a, b] \to \mathbf{C}$ be a piecewise-C^1 curve, and let ϕ be a continuous complex-valued function on γ. The Cauchy integral of ϕ over γ is the function f defined in $\mathbf{C}\backslash\gamma$ by

$$f(z) = \int_\gamma \frac{\phi(\zeta)}{\zeta - z} d\zeta.$$

Cauchy's formula for a circle thus states that the Cauchy integral of a holomorphic function over a counterclockwise oriented circle is, inside the circle, equal to $2\pi i$ times the function itself, provided the circle and its interior both lie in the open set where the function is holomorphic.

We shall prove that any Cauchy integral, and hence any holomorphic function, has local power series representations. We retain the notations above. Fix a point z_0 off γ, and let R denote the distance from z_0 to γ. Let z be a point satisfying $|z - z_0| < R$, regarded for the moment as fixed. We write $\dfrac{1}{\zeta - z}$ as a geometric series:

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - z_0) - (z - z_0)} = \frac{1}{\zeta - z_0}\left[\frac{1}{1 - \frac{z-z_0}{\zeta-z_0}}\right] = \sum_{n=0}^\infty \frac{(z-z_0)^n}{(\zeta - z_0)^{n+1}}.$$

The series converges locally uniformly (with respect to ζ) in the region $\left|\dfrac{z - z_0}{\zeta - z_0}\right| < 1$, and hence it converges uniformly on γ (a compact subset of the preceding region). The function ϕ is bounded on γ, being continuous there, and therefore the series

$$\sum_{n=0}^\infty \frac{\phi(\zeta)(z-z_0)^n}{(\zeta - z_0)^{n+1}}$$

converges uniformly on γ to $\dfrac{\phi(\zeta)}{\zeta - z}$. By VI.11 we can thus integrate the preceding series term-by-term over γ to obtain

$$f(z) = \sum_{n=0}^\infty \int_\gamma \frac{\phi(\zeta)(z-z_0)^n}{(\zeta - z_0)^{n+1}} d\zeta, \quad |z - z_0| < R.$$

In other words, letting

$$a_n = \int_\gamma \frac{\phi(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta, \quad n = 0, 1, 2, \ldots,$$

we have shown that the power series $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ represents f in the disk $|z - z_0| < R$.

As shown in V.15, a function represented by a convergent power series is differentiable to all orders, and the series is the Taylor series of the function it represents. We thus have $a_n = \dfrac{f^{(n)}(z_0)}{n!}$, which in combination with the expression above defining a_n gives an integral representation for $f^{(n)}$. Since z_0 is an arbitrary point off γ, we can write the representation as

$$f^{(n)}(z) = n! \int_\gamma \frac{\phi(\zeta)}{(\zeta - z)^{n+1}} d\zeta, \quad z \in \mathbf{C} \backslash \gamma,$$

which holds for any nonnegative integer n. The upshot is that it is legitimate to differentiate a Cauchy integral under the integral sign any number of times.

Exercise VII.7.1. Find explicitly the Cauchy integral of the constant function 1 over the interval $[0, 1]$.

VII.8. Implications for Holomorphic Functions

Let f be a holomorphic function in the open set G, and let z_0 be a point of G. For $0 < r < \operatorname{dist}(z_0, \mathbf{C} \backslash G)$, let C_r denote the circle with center z_0 and radius r, oriented counterclockwise. Then, inside C_r, the function f coincides with the Cauchy integral of $\dfrac{1}{2\pi i} f$ over C_r. This in combination with what is proved in the preceding section enables us to draw the following conclusions.

(i) *The function f is differentiable to all orders. For any positive integer n,*

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_{C_r} \frac{f(\zeta)}{(\zeta - z)^{n+1}} d\zeta, \quad |z - z_0| < r < \operatorname{dist}(z_0, \mathbf{C} \backslash G).$$

(ii) *In the disk $|z - z_0| < \operatorname{dist}(z_0, \mathbf{C} \backslash G)$, the function f is represented by its Taylor series centered at z_0.*

Exercise VII.8.1. Evaluate the following integrals, where C is the unit circle with the counterclockwise orientation.

$$(a) \int_C \frac{\sin z}{z^{38}} dz, \qquad (b) \int_C \left(\frac{z - 2}{2z - 1} \right)^3 dz.$$

Exercise VII.8.2. Evaluate

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{1-2r\cos\theta+r^2}d\theta$$

for $0<r<1$ by writing $\cos\theta=\frac{1}{2}(e^{i\theta}+e^{-i\theta})$ and reducing the given integral to a complex integral over the unit circle.

Exercise VII.8.3. Let a and b be complex numbers such that $|a|<1<|b|$. Let C be the unit circle with the counterclockwise orientation. Evaluate

$$\frac{1}{2\pi i}\int_C\frac{(z-b)^m}{(z-a)^n}dz$$

for all pairs of integers m and n.

VII.9. Cauchy Product

Once one knows that holomorphic functions are represented locally by Taylor series, it is possible to give a simplified approach to the Cauchy product of power series. Recall from Section V.17 that the Cauchy product of the two power series $\sum_{n=0}^{\infty}a_n(z-z_0)^n$ and $\sum_{n=0}^{\infty}b_n(z-z_0)^n$ is the series $\sum_{n=0}^{\infty}c_n(z-z_0)^n$ whose n-th coefficient is given by $c_n=\sum_{k=0}^{n}a_kb_{n-k}$. In Section V.17 it is proved in a direct way that if the first two series have positive radii of convergence R_1 and R_2, respectively, then their Cauchy product converges in the disk $|z-z_0|<\min\{R_1,R_2\}$ to the product of the functions they represent. We can now obtain the same conclusion more simply by an indirect argument.

To do this, let f and g denote the functions represented by the two original series in the disks $|z-z_0|<R_1$ and $|z-z_0|<R_2$, respectively. Those series are the Taylor series of the functions they represent, so that $a_n=\dfrac{f^{(n)}(z_0)}{n!}$ and $b_n=\dfrac{g^{(n)}(z_0)}{n!}$ for each n. Define the function h in the disk $|z-z_0|<\min\{R_1,R_2\}$ by $h(z)=f(z)g(z)$. Then h is holomorphic and so is represented in the preceding disk by its Taylor series centered at z_0. To verify that the last series is in fact the Cauchy product of the two original series amounts to verifying the identity

$$\frac{h^{(n)}(z_0)}{n!}=\sum_{k=0}^{n}\frac{f^{(k)}(z_0)}{k!}\cdot\frac{g^{(n-k)}(z_0)}{(n-k)!},\qquad n=0,1,2,\dots.$$

But this is just the well-known Leibniz formula for the n-th derivative of the product of two functions, which the reader has probably encountered in real analysis. The case $n=1$ is just the product rule for differentiation, and a straightforward induction argument yields the general case.

Exercise VII.9.1. If the function f is holomorphic in a neighborhood of the origin and $\sum_{n=0}^{\infty} a_n z^n$ is its Taylor series about the origin, what is the Taylor series about the origin of the function $(1-z)^{-1}f(z)$?

Exercise VII.9.2. Prove that the Taylor series about the origin of the function $\left[\mathrm{Log}(1-z)\right]^2$ is

$$\sum_{n=1}^{\infty} \frac{2H_n}{n+1} z^{n+1},$$

where $H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$ (the n-th partial sum of the harmonic series).

VII.10. Converse of Goursat's Lemma

Let f be a continuous complex-valued function in the open subset G of $\mathbf{C}$. If $\int_T f(z)dz = 0$ for every triangle T contained with its interior in G, then f is holomorphic.

Because being holomorphic is a local property, it is enough to prove this for the case where G is a disk. In that case, as pointed out in Section VII.3, the hypotheses imply that f has a primitive, in other words, there is a holomorphic function g in G such that $g' = f$. Since holomorphic functions have holomorphic derivatives, the desired conclusion follows.

R. B. Burckel, on page 188 of his book cited in the reference list, states that the preceding result "is (essentially) contained" in an 1886 paper of G. Morera. In a letter to the author, Burckel explained that it is hard to pin down what Morera actually proves because of imprecise language, but what he seems to assume is that the function in question has integral 0 around *all* closed curves in its region. That that assumption implies holomorphicity is often referred to as Morera's theorem.

The first exercise below gives a stronger Morera-type theorem than the one involving triangles, stronger in the sense that the one involving triangles is easily deduced from it.

Exercise VII.10.1. Prove that if f is a continuous complex-valued function in the open subset G of $\mathbf{C}$, and if $\int_R f(z)dz = 0$ for every rectangle R, with edges parallel to the coordinate axes, contained with its interior in G, then f is holomorphic.

Exercise VII.10.2. (A version of the Schwarz reflection principle.) Let the function f be continuous in the region $\{z : |z| < 1,\ \mathrm{Im}\, z \geq 0\}$, real valued on the segment $(-1, 1)$ of the real axis, and holomorphic in the open set $\{z : |z| < 1, \mathrm{Im}\, z > 0\}$. Use the result of the preceding exercise to prove f

can be extended holomorphically to the open unit disk. (Suggestion: See Exercise II.8.2.)

VII.11. Liouville's Theorem

A function that is holomorphic in all of $\mathbf{C}$ is called an entire function. The theorem of Liouville states that *the only bounded entire functions are the constant functions.*

For the proof, assume f is a bounded entire function, say $|f(z)| \le M$ for all z. Fix a point z_0 in $\mathbf{C}$, and for $R > 0$ let C_R denote the circle $|z - z_0| = R$, with the counterclockwise orientation. By VII.8(i) we have

$$f'(z_0) = \frac{1}{2\pi i} \int_{C_R} \frac{f(z)}{(z - z_0)^2} dz.$$

The integrand in the preceding integral is bounded in absolute value by $\frac{M}{R^2}$, giving the estimate

$$|f'(z_0)| \le \frac{M}{2\pi R^2} L\Big(C_R\Big) = \frac{M}{R}.$$

This being true for all $R > 0$, we have $f'(z_0) = 0$. Since the derivative of f vanishes identically, the function f must be constant, as desired.

Exercise VII.11.1. Let f be an entire function such that, for some positive integer n and positive number R, $|f(z)/z^n|$ is bounded for $|z| > R$. Prove f is a polynomial of degree at most n.

Exercise* VII.11.2. Let f be a holomorphic map of the open unit disk into itself. Prove that

$$|f'(z)| \le \frac{1}{1 - |z|}$$

for all z in the disk.

Exercise VII.11.3. Prove that an entire function with a positive real part is constant.

VII.12. Fundamental Theorem of Algebra

Every nonconstant polynomial with complex coefficients can be factored over $\mathbf{C}$ *into linear factors.*

Purely algebraic reasoning reduces this to the statement that every nonconstant polynomial with complex coefficients has a complex root. In fact, suppose the preceding statement is known, and let p be a nonconstant polynomial with complex coefficients. Let z_0 be a root of p. The standard division algorithm of algebra then enables one to write $p(z) = (z - z_0)q(z)$, where q is a polynomial whose degree is one less than that of p. The same reasoning can now be applied to q, and so on, yielding the desired linear factorization in finitely many steps.

To prove that nonconstant polynomials with complex coefficients have complex roots we argue by contradiction. Suppose $p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ is a polynomial of positive degree n, with complex coefficients, but without complex roots. Let $f = \frac{1}{p}$. Then f is an entire function. Since $a_n \neq 0$ we can write, for $z \neq 0$,

$$p(z) = a_n z^n \left(1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_1}{a_n z^{n-1}} + \frac{a_0}{a_n z^n}\right).$$

Consequently,

$$|p(z)| \geq |a_n z^n| \left(1 - \frac{|a_{n-1}|}{|a_n z|} - \cdots - \frac{|a_1|}{|a_n z^{n-1}|} - \frac{|a_0|}{|a_n z^n|}\right).$$

The quantity in parentheses on the right will be no smaller than $\frac{1}{2}$ when $|z|$ is sufficiently large, say when $|z| \geq R$. Hence

$$|f(z)| \leq \frac{2}{|a_n| R^n}, \qquad |z| \geq R.$$

On the other hand, since $|f|$ is a continuous function, it is bounded on the compact set $|z| \leq R$. Thus f is a bounded entire function, and so it is constant by Liouville's theorem. This is the desired contradiction because the polynomial p was assumed to be nonconstant.

The proof just given is purely an existence proof; it does not provide one with a method of locating the roots of a polynomial. The problem of finding accurate approximations to the roots of complex polynomials is a major theme in numerical analysis.

VII.13. Zeros of Holomorphic Functions

Zeros of holomorphic functions can be treated like zeros of polynomials. Suppose f is a holomorphic function in the open subset G of $\mathbf{C}$, and suppose z_0 is a point of G such that $f(z_0) = 0$. The point z_0 is called a zero of f of order (or multiplicity) m, where m is a positive integer, if $f^{(n)}(z_0) = 0$ for $n = 0, \ldots, m-1$, while $f^{(m)}(z_0) \neq 0$. In that case the Taylor series for f about z_0 has the form $\sum_{n=m}^{\infty} a_n (z - z_0)^n$, where $a_m = \frac{f^{(m)}(z_0)}{m!} \neq 0$. We can write $f(z) = (z - z_0)^m g(z)$, where the function g, defined by

$$g(z) = \begin{cases} (z - z_0)^{-m} f(z), & z \in G \setminus \{z_0\} \\ a_m, & z = z_0, \end{cases}$$

is holomorphic in G. The function g, being nonzero at z_0, is nonzero in some neighborhood of z_0, implying that z_0 is at a positive distance from $f^{-1}(0)\backslash\{z_0\}$, the set of all the other zeros of f. In other words, zeros of finite order of holomorphic functions are isolated.

The point z_0 is called a simple zero of f if it is a zero of order 1. In that case the function g defined above takes the value $f'(z_0)$ at z_0. The reasoning above shows that if z_0 is a zero of f of any order, then the function that equals $(z-z_0)^{-1}f(z)$ in $G\backslash\{z_0\}$ and equals $f'(z_0)$ at z_0 is holomorphic.

What about zeros of infinite order, that is, points at which f and all of its derivatives vanish? It turns out that these exist only in trivial cases. First, if G is connected, then f can have a zero of infinite order only if it is the zero function. In fact, the zeros of f of infinite order form an open set (since the Taylor series of the function centered at such a zero has all of its coefficients equal to 0), and, by the discussion in the preceding paragraph, so do the points of G that are not zeros of f of infinite order. If G is connected, then one of these sets must be empty. If G is not connected, and if f has a zero of infinite order at the point z_0 of G, one can at least conclude that f vanishes identically in the connected component of G containing z_0.

In contrast to the behavior of holomorphic functions, it is possible for a real-valued C^∞ function on $\mathbf{R}$ to have a zero of infinite order and yet not vanish identically. A standard example is the function that equals 0 at the origin and equals e^{-1/x^2} elsewhere. One can show that the origin is a zero of infinite order of this function.

Most often in complex function theory one is concerned with holomorphic functions in connected open sets. The following statement summarizes what is established above about the zeros of such functions. *Let f be a holomorphic function, not the zero function, in the connected open subset G of $\mathbf{C}$. Then each zero of f is of finite order, and $f^{-1}(0)$, the zero set of f, has no limit points in G.* Since an uncountable subset of $\mathbf{C}$ always contains limit points of itself, one can conclude in particular that the set $f^{-1}(0)$ is at most countable. The same is true of $f^{-1}(w)$, for any complex number w, provided f is not the constant function w.

Exercise VII.13.1. Prove that there is no holomorphic function f in the open unit disk such that $f(\frac{1}{n}) = 2^{-n}$ for $n = 2, 3, \ldots$.

Exercise VII.13.2. Let f be a holomorphic function in the open subset G of $\mathbf{C}$. Let the point z_0 of G be a zero of f of order m. Prove that there is a branch of $f^{\frac{1}{m}}$ in some open disk centered at z_0.

VII.14. The Identity Theorem

Let f and g be holomorphic functions in the connected open subset G of $\mathbf{C}$. If $f(z) = g(z)$ for all z in a subset of G that has a limit point in G, then $f = g$.

Thus, for example, a holomorphic function in the open unit disk is uniquely determined by the values it assumes on the set $\{2^{-n} : n = 1, 2, 3, \dots\}$.

To prove the theorem one notes that, if f and g agree on a subset of G that has a limit point in G, then the zero set of the function $f - g$ has a limit point in G, which implies by what is proved in the last section that $f - g$ is the zero function.

Exercise VII.14.1. Prove that there is no holomorphic function f in the open unit disk such that $f\left(\frac{1}{n}\right) = \frac{(-1)^n}{n^2}$ for $n = 2, 3, \dots$.

Exercise VII.14.2. (a) Let G be a nonempty, connected, open subset of $\mathbf{C}$ which is symmetric with respect to the real axis. Let f be a holomorphic function in G such that f is real-valued on $G \cap \mathbf{R}$. Prove that $f(\overline{z}) = \overline{f(z)}$ for all z in G. (Suggestion: Use Exercise II.8.2.)

(b) Let G be as in part (a), and let f be a holomorphic function in G such that f is real-valued on a nonempty subinterval of $G \cap \mathbf{R}$. Prove that f is real-valued on all of $G \cap \mathbf{R}$.

VII.15. Weierstrass Convergence Theorem

Let G be an open subset of $\mathbf{C}$ and $(f_k)_{k=1}^{\infty}$ a sequence of holomorphic functions in G that converges locally uniformly in G to the function f. Then f is holomorphic, and for each positive integer n, the sequence $\left(f_k^{(n)}\right)_{k=1}^{\infty}$ converges locally uniformly in G to $f^{(n)}$.

Once again we have a contrast between real function theory and complex function theory. The local uniform convergence on $\mathbf{R}$ of a sequence of real-valued C^∞ functions does not even imply the differentiability, let alone the infinite differentiability, of the limit function. In fact, by the classical Weierstrass approximation theorem, any continuous real-valued function on $\mathbf{R}$, even a nowhere differentiable one, is the local uniform limit of a sequence of polynomials.

To prove the Weierstrass convergence theorem, fix a point z_0 in G, and let r be a positive number such that the disk $|z - z_0| \le r$ is contained in G. Let C_r denote the circle $|z - z_0| = r$, with the counterclockwise orientation.

For each positive integer k we have, by Cauchy's formula for a circle,

$$f_k(z) = \frac{1}{2\pi i}\int_{C_r} \frac{f_k(\zeta)}{\zeta - z}d\zeta, \quad |z - z_0| < r.$$

As $k \to \infty$ we have $f_k \to f$ uniformly on C_r, so, for fixed z satisfying $|z - z_0| < r$, the integrand in the preceding integral converges uniformly on C_r to $\dfrac{f(\zeta)}{\zeta - z}$. Therefore, by VI.11, the integral converges to $\displaystyle\int_{C_r} \frac{f(\zeta)}{\zeta - z}d\zeta$. Since also $f_k(z) \to f(z)$, we obtain

$$f(z) = \frac{1}{2\pi i}\int_{C_r} \frac{f(\zeta)}{\zeta - z}d\zeta, \quad |z - z_0| < r.$$

Thus f is given in the disk $|z - z_0| < r$ by a Cauchy integral, so it is holomorphic in that disk. Since G is covered by such disks, f is holomorphic in G.

Now fix a positive integer n. By VII.8(i) we have

$$f^{(n)}(z) = \frac{n!}{2\pi i}\int_{C_r} \frac{f(\zeta)}{(\zeta - z)^{n+1}}d\zeta, \quad |z - z_0| < r,$$

and similarly for each function f_k. Hence

$$f^{(n)}(z) - f_k^{(n)}(z) = \frac{n!}{2\pi i}\int_{C_r} \frac{f(\zeta) - f_k(\zeta)}{(\zeta - z)^{n+1}}d\zeta, \quad |z - z_0| < r.$$

Let M_k denote the maximum of $|f(\zeta) - f_k(\zeta)|$ for ζ on C_r, $k = 1, 2, \ldots$. Then, for $|z - z_0| < \frac{r}{2}$, the absolute value of the integrand in the preceding integral is bounded by $\dfrac{2^{n+1}M_k}{r^{n+1}}$. The usual estimate of the integral therefore gives us

$$\begin{aligned} \left|f^{(n)}(z) - f_k^{(n)}(z)\right| &\leq \frac{n!}{2\pi} \cdot \frac{2^{n+1}M_k}{r^{n+1}} \cdot L(C_r) \\ &= \frac{2^{n+1}n!M_k}{r^n}, \qquad |z - z_0| < \frac{r}{2}. \end{aligned}$$

Since $M_k \to 0$, it follows that the sequence $\left(f_k^{(n)}\right)_{k=1}^{\infty}$ converges uniformly to $f^{(n)}$ in the disk $|z - z_0| < \frac{r}{2}$. This establishes the local uniform convergence of the sequence $\left(f_k^{(n)}\right)_{k=1}^{\infty}$ to $f^{(n)}$, completing the proof of the theorem.

VII.16. Maximum Modulus Principle

Let f be a nonconstant holomorphic function in the open connected subset G of $\mathbf{C}$. Then $|f|$ does not attain a local maximum in G.

An immediate consequence is that, if K is a compact subset of G, then $|f|$ attains its maximum over K only at points of the boundary of K. This is the form in which the principle is usually applied.

To establish the maximum modulus principle, we argue by contradiction, assuming that there is a point z_0 of G at which $|f|$ attains a local maximum. Choose a positive number r such that the disk $|z - z_0| \leq r$ is contained in G, and such that $|f(z)| \leq |f(z_0)|$ for all points z in that disk. If $f(z_0) = 0$, then it follows immediately from the identity theorem that f is the zero function, contrary to the hypothesis that f is nonconstant. We thus may as well assume that $f(z_0) \neq 0$. Let $\lambda = \dfrac{|f(z_0)|}{f(z_0)}$. By the mean value property,

$$f(z_0) \;=\; \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{it})dt.$$

Hence

$$\begin{aligned}
|f(z_0)| = \lambda f(z_0) \;&=\; \frac{1}{2\pi}\int_0^{2\pi} \lambda f(z_0 + re^{it})dt \\
&= \mathrm{Re}\,\frac{1}{2\pi}\int_0^{2\pi} \lambda f(z_0 + re^{it})dt \\
&= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{Re}\Big[\lambda f(z_0 + re^{it})\Big]dt,
\end{aligned}$$

which we can rewrite as

$$\frac{1}{2\pi}\int_0^{2\pi}\Big(|f(z_0)| - \mathrm{Re}\big[\lambda f(z_0 + re^{it})\big]\Big)dt \;=\; 0.$$

The integrand in the preceding integral is continuous and nonnegative, so the vanishing of the integral implies the vanishing of the integrand. Thus $\mathrm{Re}\,\lambda f(z) = |f(z_0)|$ for z on C_r. This together with the inequality $|\lambda f(z)| \leq |f(z_0)|$ implies $\lambda f(z) = |f(z_0)|$, in other words, $f(z) = f(z_0)$, for z on C_r. Thus, by the identity theorem, f is the constant function $f(z_0)$, contrary to the hypothesis that f is nonconstant. This completes the proof.

Exercise VII.16.1. Let f be a nonconstant holomorphic function in the connected open subset G of $\mathbf{C}$. Prove that $|f|$ can attain a local minimum in G only at a zero of f.

Exercise VII.16.2. Give a proof of the fundamental theorem of algebra based on the maximum modulus principle.

VII.17. Schwarz's Lemma

Let f be a holomorphic map of the open unit disk into itself such that $f(0) = 0$. Then $|f(z)| \leq |z|$ for all z in the disk. The inequality is strict at all points *other than the origin, except in the case where f has the form $f(z) = \lambda z$ with λ a constant of unit modulus.*

To establish this we define the function g in the disk by

$$g(z) = \begin{cases} \dfrac{f(z)}{z}, & 0 < |z| < 1 \\ f'(0), & z = 0. \end{cases}$$

Then g is holomorphic (see Section VII.13). For $0 < r < 1$ the function g is bounded in absolute value by $\frac{1}{r}$ on the circle $|z| = r$, and hence it has the same bound in the disk $|z| \leq r$, by the maximum modulus principle. This being true for all r in $(0, 1)$, the function g is bounded in absolute value by 1, which means $|f(z)| \leq |z|$, as desired. Moreover, if the preceding inequality is an equality for any z other than the origin, then $|g|$ has a local maximum at that z. Then, again by the maximum modulus principle, g is a constant, λ, which means $f(z) = \lambda z$.

Exercise* VII.17.1. Prove, under the hypotheses of Schwarz's lemma, that $|f'(0)| \leq 1$, and that the inequality is strict unless $f(z) = \lambda z$ with λ a constant of unit modulus.

Exercise VII.17.2. (Pick's lemma.) Let f be a holomorphic map of the open unit disk into itself. Prove that, for any two points z and w in the disk,

$$\left| \frac{f(z) - f(w)}{1 - f(z)\overline{f(w)}} \right| \leq \left| \frac{z - w}{1 - z\overline{w}} \right|,$$

and that the inequality is strict for $z \neq w$ except when f is a linear-fractional transformation mapping the disk onto itself.

Exercise VII.17.3. Let f be a holomorphic map of the unit disk into itself. Prove that

$$|f'(w)| \leq \frac{1 - |f(w)|^2}{1 - |w|^2}$$

for all w in the disk, and that the inequality is strict except when f is a linear fractional transformation mapping the disk onto itself. Compare with Exercise VII.11.2.

Exercise VII.17.4. Let f be a holomorphic map from the half-plane $\operatorname{Re} z > 0$ into the open unit disk. Prove that

$$|f'(z)| \leq \frac{1 - |f(z)|^2}{2 \operatorname{Re} z}$$

for all z in the half-plane.

Exercise VII.17.5. Prove that, except for the identity function, a holomorphic map of the open unit disk into itself has at most one fixed point in the disk.

VII.18. Existence of Harmonic Conjugates

Let u be a real-valued harmonic function in the convex open subset G of $\mathbf{C}$. *Then there is a holomorphic function g in G such that $u = \operatorname{Re} g$. The function g is unique to within addition of an imaginary constant.*

In other words, u has a harmonic conjugate in G, unique to within addition of a constant. The conclusion is true under a much weaker assumption on G than convexity, as will be shown later on the basis of the general Cauchy theorem.

To prove the proposition we consider the function $f = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y} \left(= 2\frac{\partial u}{\partial z}\right)$, which is holomorphic in G (by Exercise II.16.2). As shown in the proof of Cauchy's theorem for a convex region, the function f has a primitive in G; that is, there is a holomorphic function g in G such that $g' = f$. Fix a point z_0 in G. By adding a constant to g, we may assume $g(z_0) = u(z_0)$. Let $u_1 = \operatorname{Re} g$ and $v_1 = \operatorname{Im} g$. We have

$$g' = \frac{\partial u_1}{\partial x} + i\frac{\partial v_1}{\partial x} = \frac{\partial v_1}{\partial y} - i\frac{\partial u_1}{\partial y},$$

and also

$$g' = f = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}.$$

Hence $\frac{\partial u}{\partial x} = \frac{\partial u_1}{\partial x}$ and $\frac{\partial u}{\partial y} = \frac{\partial u_1}{\partial y}$. The first partial derivatives of the function $u - u_1$ thus vanish throughout G, implying that $u - u_1$ is constant. Since $u(z_0) = u_1(z_0)$, the constant is 0, and $u = \operatorname{Re} g$, as desired.

As for uniqueness, if two holomorphic functions in G have the same real part, then their difference has derivative 0, by the Cauchy-Riemann equations. In that case the functions differ by a constant.

Exercise VII.18.1. Prove that a real-valued harmonic function u in the open unit disk has a series representation

$$u(re^{i\theta}) = a_0 + \sum_{n=1}^{\infty} r^n (a_n \cos n\theta + b_n \sin n\theta),$$

where the coefficients a_n and b_n are real numbers.

Exercise VII.18.2. Prove that a positive harmonic function in $\mathbf{C}$ is constant.

Exercise VII.18.3. Let u be a nonconstant real-valued harmonic function in $\mathbf{C}$.

Prove that the set $u^{-1}(c)$ is unbounded for every real number c.

VII.19. Infinite Differentiability of Harmonic Functions

Harmonic functions are of class C^∞.

Because differentiability is a local property, it will suffice to prove this for a harmonic function u in an open disk D. We may assume without loss of generality that u is real-valued. Then, by VII.18, we can write $u = \operatorname{Re} g$ with g holomorphic in D. The desired conclusion now follows by the infinite differentiability of holomorphic functions.

VII.20. Mean Value Property for Harmonic Functions

Let u be a harmonic function in the open subset G of $\mathbf{C}$, and let z_0 be a point of G. Then

$$u(z_0) = \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + re^{it})dt, \quad 0 < r < \operatorname{dist}(z_0, \mathbf{C}\backslash G).$$

It will be enough to prove this for the case where u is real-valued. In that case there is, by VII.18, a holomorphic function in the disk $|z - z_0| < \operatorname{dist}(z_0, \mathbf{C}\backslash G)$ whose real part is u. The mean value property for u then follows by VII.6, the corresponding property for holomorphic functions.

VII.21. Identity Theorem for Harmonic Functions

If G is a connected open subset of $\mathbf{C}$, and if u and v are harmonic functions in G that agree on a nonempty open subset of G, then $u = v$ throughout G.

It will be enough to prove this for the case where u is real-valued and v is the zero function. Suppose then that u is a real-valued harmonic function in G that vanishes on a nonempty open subset of G. Let H be the interior of the set $u^{-1}(0)$. Then H is nonempty and open. Suppose H is not all of G. Then, because G is connected, some point z_0 of G lies in the boundary of H. Choose $r > 0$ such that the disk $|z - z_0| < r$ is contained in G. By VII.18 there is a holomorphic function g in this disk such that $u = \operatorname{Re} g$ and $g(z_0) = u(z_0)$ $(= 0)$. Since z_0 is in the boundary of H, the intersection of H with the disk $|z - z_0| < r$ is nonempty. Thus $\operatorname{Re} g$ vanishes on a nonempty subdisk of the preceding disk. On the smaller disk we have $g' = 0$ by the Cauchy-Riemann equations. Hence $g' = 0$ in the disk $|z - z_0| < r$, by the identity theorem for holomorphic functions. Therefore g is constant in the disk $|z - z_0| < r$. Since $g(z_0) = 0$, the constant is 0. We have thus shown that u vanishes in the disk $|z - z_0| < r$, which means z_0 belongs to the set H (the interior of $u^{-1}(0)$). This is a contradiction because z_0 was chosen to be in the boundary of H. We can conclude that $H = G$, in other words, u is the zero function, as desired.

VII.22. Maximum Principle for Harmonic Functions

Let u be a nonconstant real-valued harmonic function in the connected open subset G of $\mathbf{C}$. Then u does not attain a local maximum in G.

This is proved from the mean value property and the identity theorem in the same way that the corresponding result for holomorphic functions is proved in Section VII.16. Of course, by applying the maximum principle to $-u$, we see that u also does not attain a local minimum in G.

VII.23. Harmonic Functions in Higher Dimensions

The properties given in Sections VII.19 through VII.22 hold also for harmonic functions of more than two variables. The proofs given here, however, since they are based on complex function theory, are not available in higher dimensions. The reader wishing to pursue harmonic functions in greater depth can consult one of the books on the subject, for example, the one by S. Axler, P. Bourdon and W. Ramey in the list of references.

Chapter VIII

Laurent Series and Isolated Singularities

Laurent series are power series that include negative as well as positive powers of the variable. Such series arise, in particular, when one examines holomorphic functions near what are called isolated singularities.

VIII.1. Simple Examples

We have seen that the function $\dfrac{1}{1-z}$ is represented in the disk $|z| < 1$ by the geometric series $\sum_{n=0}^{\infty} z^n$. By writing $\dfrac{1}{1-z} = -\dfrac{1}{z}\left(\dfrac{1}{1-\frac{1}{z}}\right)$, we see that $\dfrac{1}{1-z}$ is represented in the region $|z| > 1$ by a different series, the series $-\sum_{n=1}^{\infty} z^{-n}$, a series in negative powers of z.

Consider the function $\dfrac{1}{(z-1)(z-2)}$, for a slightly more complicated example, which can be written as $\dfrac{1}{z-2} - \dfrac{1}{z-1}$. The function $\dfrac{1}{z-2}$ is represented in the disk $|z| < 2$ by a power series in z, and we have just seen that the function $\dfrac{1}{1-z}$ is represented in the region $|z| > 1$ by a series in negative powers of z. Hence, by using a series in both positive and negative powers of z, we can represent the function $\dfrac{1}{(z-1)(z-2)}$ in the annulus $1 < |z| < 2$.

Exercise VIII.1.1. Let a and b be complex numbers such that $0 < |a| < |b|$. Find a series in positive and negative powers of z that represents the function $\dfrac{1}{(z-a)(z-b)}$ in the annulus $|a| < |z| < |b|$.

VIII.2. Laurent Series

A Laurent series centered at the point z_0 of $\mathbf{C}$ is, by definition, a series of the form $\sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$, where the coefficients a_n are complex constants. The series is said to converge at the point z if both of the series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ and $\sum_{n=1}^{\infty} a_{-n}(z-z_0)^{-n}$ converge. In that case

$$\lim_{N\to\infty} \sum_{n=-N}^{N} a_n(z-z_0)^n$$

exists and is defined to be the sum of the series.

Let

$$\begin{aligned} R_1 &= \limsup_{n\to\infty} |a_{-n}|^{1/n}, \\ R_2 &= \frac{1}{\limsup\limits_{n\to\infty} |a_n|^{1/n}}. \end{aligned}$$

By the Cauchy-Hadamard theorem V.12, the series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ converges absolutely and locally uniformly in the disk $|z-z_0| < R_2$, and the series $\sum_{n=1}^{\infty} a_{-n}(z-z_0)^{-n}$ converges absolutely and locally uniformly in the region $|z-z_0| > R_1$. Hence, if $R_1 < R_2$, then the Laurent series $\sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$ converges absolutely and locally uniformly in the annulus $R_1 < |z-z_0| < R_2$, called the annulus of convergence of the series.

Assume $R_1 < R_2$, and let f be the function to which the Laurent series converges in the annulus $R_1 < |z-z_0| < R_2$, in other words, let $f(z) = \lim\limits_{N\to\infty} \sum_{n=-N}^{N} a_n(z-z_0)^n$. Then f is holomorphic, by the Weierstrass convergence theorem VII.15. For $R_1 < r < R_2$, let C_r denote the circle $|z-z_0| = r$, oriented counterclockwise. The series then converges uniformly to f on C_r, enabling us to evaluate the integral $\int_{C_r} f(z)dz$ by integrating the series term-by-term. We obtain

$$\begin{aligned} \int_{C_r} f(z)dz &= \sum_{n=-\infty}^{\infty} a_n \int_{C_r} (z-z_0)^n dz \\ &= 2\pi i a_{-1}. \end{aligned}$$

Thus

$$a_{-1} = \frac{1}{2\pi i} \int_{C_r} f(z)dz.$$

Applying this to the function $(z-z_0)^{-k-1}f(z)$ in place of f, where k is any integer, we find that

$$a_k = \frac{1}{2\pi i}\int_{C_r}(z-z_0)^{-k-1}f(z)dz.$$

The coefficients of the series are thus uniquely determined by the function to which the series converges.

Exercise VIII.2.1. Determine the annulus of convergence of the Laurent series $\sum_{n=-\infty}^{\infty} a^{n^2}z^n$, where $0 < |a| < 1$.

VIII.3. Cauchy Integral Near ∞

We shall show that any function given by a Cauchy integral can be represented "near ∞" by a Laurent series in negative powers of $z - z_0$. The reasoning is essentially the same as that used in Section VII.7 to obtain power series representations of Cauchy integrals. The result to be established will be used later to show that any function holomorphic in an annulus has a Laurent series representation.

Let $\gamma : [a,b] \to \mathbf{C}$ be a piecewise-C^1 curve, and let ϕ be a continuous complex-valued function defined on γ. Let f be the Cauchy integral of ϕ over γ, that is, the function defined in $\mathbf{C}\backslash\gamma$ by

$$f(z) = \int_\gamma \frac{\phi(\zeta)}{\zeta - z}d\zeta.$$

Fix a point z_0 in $\mathbf{C}$, and let R denote the maximum of $|\zeta - z_0|$ for ζ on γ. Fix any point z such that $|z-z_0| > R$. We have

$$\begin{aligned}\frac{1}{\zeta - z} &= \frac{1}{(\zeta - z_0) - (z - z_0)}\\ &= \frac{-1}{(z-z_0)}\left[\frac{1}{1-\frac{\zeta - z_0}{z-z_0}}\right]\\ &= -\sum_{n=1}^{\infty}(\zeta - z_0)^{n-1}(z-z_0)^{-n},\end{aligned}$$

the series converging locally uniformly (with respect to ζ) in the region $|\zeta - z_0| < |z - z_0|$, and hence uniformly on γ. The series

$$-\sum_{n=1}^{\infty}\phi(\zeta)(\zeta - z_0)^{n-1}(z-z_0)^{-n}$$

thus converges uniformly to $\dfrac{\phi(\zeta)}{\zeta - z}$ on γ. We can therefore integrate the series term-by-term over γ to obtain

$$f(z) = -\sum_{n=1}^{\infty}(z-z_0)^{-n}\int_\gamma (\zeta - z_0)^{n-1}\phi(\zeta)d\zeta\,;$$

in other words, the Laurent expansion $f(z) = \sum_{n=-\infty}^{-1} a_n(z-z_0)^n$, with

$$a_n = -\int_\gamma (\zeta - z_0)^{-n-1}\phi(\zeta)d\zeta,$$

holds in the region $|z - z_0| > R$.

VIII.4. Cauchy's Theorem for Two Concentric Circles

Let the function f be holomorphic in the annulus $R_1 < |z - z_0| < R_2$. For $R_1 < r < R_2$, let C_r denote the circle $|z - z_0| = r$, oriented counterclockwise. Then $\int_{C_r} f(z)dz$ is independent of r ($R_1 < r < R_2$).

To prove this, let $I(r)$ denote the preceding integral. Introducing the parametrization $t \to z_0 + re^{it}$ ($0 \le t \le 2\pi$) of C_r, we can write

$$I(r) = \int_0^{2\pi} f(z_0 + re^{it})ire^{it}dt.$$

The integrand in the preceding integral, as a function of the two real variables r and t, has continuous partial derivatives. By a standard theorem in calculus (whose proof is in Appendix 4), we can find $\dfrac{dI(r)}{dr}$ by differentiating under the integral sign:

$$\frac{dI(r)}{dr} = \int_0^{2\pi}\frac{\partial}{\partial r}\Big[f\big(z_0 + re^{it}\big)ire^{it}\Big]dt.$$

We have

$$\frac{\partial}{\partial r}\Big[f\big(z_0 + re^{it}\big)ire^{it}\Big] = f'\big(z_0+re^{it}\big)ire^{2it} + f\big(z_0+re^{it}\big)ie^{it} = \frac{\partial}{\partial t}\Big[f\big(z_0 + re^{it}\big)e^{it}\Big].$$

Consequently

$$\frac{dI(r)}{dr} = \int_0^{2\pi}\frac{\partial}{\partial t}\Big[f\big(z_0 + re^{it}\big)e^{it}\Big]dt = f\big(z_0 + re^{it}\big)e^{it}\Big|_{t=0}^{t=2\pi} = 0,$$

the desired conclusion.

Exercise VIII.4.1. Let p and q be polynomials such that deg $q >$ deg p + 1. Let C be a circle whose interior contains all of the roots of q. Prove that

$$\int_C \frac{p(z)}{q(z)}dz = 0.$$

VIII.5. Cauchy's Formula for an Annulus

Let the function f be holomorphic in the annulus $R_1 < |z - z_0| < R_2$. For $R_1 < r < R_2$, let C_r denote the circle $|z-z_0| = r$, oriented counterclockwise. If $R_1 < r_1 < |w - z_0| < r_2 < R_2$, then

$$f(w) = \frac{1}{2\pi i}\int_{C_{r_2}} \frac{f(z)}{z-w}dz - \frac{1}{2\pi i}\int_{C_{r_1}} \frac{f(z)}{z-w}dz.$$

To prove this we fix w as above and define the function g in the annulus $R_1 < |z - z_0| < R_2$ by

$$g(z) = \begin{cases} \dfrac{f(z)-f(w)}{z-w}, & z \neq w \\ f'(w), & z = w. \end{cases}$$

Then g is holomorphic (see Section VII.13), so by VIII.4 we have

$$\int_{C_{r_1}} g(z)dz = \int_{C_{r_2}} g(z)dz.$$

This we can rewrite as

$$\int_{C_{r_2}} \frac{f(z)}{z-w}dz - \int_{C_{r_1}} \frac{f(z)}{z-w}dz = f(w)\int_{C_{r_2}} \frac{1}{z-w}dz - f(w)\int_{C_{r_1}} \frac{1}{z-w}dz.$$

The first integral on the right side is $2\pi i$, by Cauchy's formula for a circle, and the second integral is 0, by Cauchy's theorem for a convex region. The right side therefore equals $2\pi i f(w)$, as desired.

VIII.6. Existence of Laurent Series Representations

Let the function f, as before, be holomorphic in the annulus $R_1 < |z-z_0| < R_2$. Fix r_1 and r_2 satisfying $R_1 < r_1 < r_2 < R_2$. By VIII.5 (with slightly altered notation),

$$f(z) = \frac{1}{2\pi i}\int_{C_{r_2}} \frac{f(\zeta)}{\zeta - z}d\zeta - \frac{1}{2\pi i}\int_{C_{r_1}} \frac{f(\zeta)}{\zeta - z}d\zeta, \quad r_1 < |z - z_0| < r_2.$$

In other words, in the annulus $r_1 < |z-z_0| < r_2$ we have $f = f_2 + f_1$, where

$$f_2(z) = \frac{1}{2\pi i}\int_{C_{r_2}} \frac{f(\zeta)}{\zeta - z}d\zeta, \quad |z - z_0| < r_2;$$

$$f_1(z) = \frac{-1}{2\pi i}\int_{C_{r_1}} \frac{f(\zeta)}{\zeta - z}d\zeta, \quad |z - z_0| > r_1.$$

As shown in Section VII.7, the function f_2 has in the disk $|z - z_0| < r_2$ the power series representation $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$, where

$$a_n = \frac{1}{2\pi i} \int_{C_{r_2}} (\zeta - z_0)^{-n-1} f(\zeta) d\zeta.$$

As shown in Section VIII.3, the function f_1 has in the region $|z - z_0| > r_1$ the Laurent series representation $f(z) = \sum_{n=-\infty}^{-1} a_n(z - z_0)^n$, where

$$a_n = \frac{1}{2\pi i} \int_{C_{r_1}} (\zeta - z_0)^{-n-1} f(\zeta) d\zeta.$$

Consequently, f has in the annulus $r_1 < |z - z_0| < r_2$ the Laurent series representation $f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$. By VIII.4, if $n \geq 0$ the integral defining a_n is independent of r_2 (provided, of course, that $R_1 < r_2 < R_2$), and if $n < 0$ the integral defining a_n is independent of r_1 (with the analogous proviso). Since r_2 can be chosen arbitrarily close to R_2 and r_1 can be chosen arbitrarily close to R_1, we can conclude that the Laurent series representation $f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$ holds in the entire annulus $R_1 < |z - z_0| < R_2$.

VIII.7. Isolated Singularities

By a punctured disk in $\mathbf{C}$ centered at z_0, one means an open annulus with center z_0 and inner radius 0. Such a region is defined by a pair of inequalities, $0 < |z - z_0| < R$, where R, the radius of the punctured disk, is positive. The case $R = \infty$ is not excluded, but in that case one usually speaks of a punctured plane rather than a punctured disk.

Let f be a holomorphic function defined in the open subset G of $\mathbf{C}$. The point z_0 is called an isolated singularity of f if z_0 is not in G but G contains a punctured disk centered at z_0. Then, as shown in the preceding section, there is a Laurent series representation $f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$ converging in the punctured disk $0 < |z - z_0| < R$, where R is the distance from z_0 to $\mathbf{C} \backslash G$ (interpreted as ∞ in case $G = \mathbf{C} \backslash \{z_0\}$). The series $\sum_{n=-\infty}^{-1} a_n(z - z_0)^n$ is called the principal part of the Laurent series of f at z_0. The point z_0 is classified as either a removable singularity, a pole, or an essential singularity of f according to the following scheme.

The point z_0 is called a removable singularity of f if $a_n = 0$ for every negative integer n, in other words, if the principal part of the Laurent series of f at z_0 is trivial. In that case one can extend f holomorphically to the open set $G \cup \{z_0\}$ by defining $f(z_0) = a_0$. Henceforth, functions with removable singularities will be assumed to be so extended.

An example of a removable singularity is provided by the function $f(z) = \frac{z}{e^z - 1}$, defined originally in $\mathbf{C} \backslash 2\pi i \mathbf{Z}$. The function $e^z - 1$ that appears in the denominator is entire and has a simple zero at the origin, so it can be written as $e^z - 1 = zg(z)$, where g is entire and $g(0) \neq 0$ (in fact $g(0) = 1$). The function $\frac{1}{g}$ is thus holomorphic in a neighborhood of the origin, and it coincides with f in a punctured disk centered at the origin. It follows that the origin is a removable singularity of f.

The point z_0 is called a pole of f of order m, where m is a positive integer, if $a_{-m} \neq 0$ but $a_n = 0$ for $n < -m$. In that case the principal part of the Laurent series of f at z_0 is a rational function. An example is provided by the function $f(z) = z^{-m}$, which has a pole of order m at the origin.

If f has a pole of order m at z_0, then the function $g(z) = (z - z_0)^m f(z)$ has a removable singularity at z_0; we can extend it holomorphically to $G \cup \{z_0\}$ by defining $g(z_0) = a_{-m}$. Thus $f(z) = (z - z_0)^{-m} g(z)$ with g holomorphic in $G \cup \{z_0\}$ and nonzero at z_0. It follows that the function $\frac{1}{f}$ has a removable singularity at z_0; its holomorphic extension to a neighborhood of z_0 has a zero of order m at z_0. In particular, if z_0 is a pole of f, then $\lim_{z \to z_0} f(z) = \infty$.

A pole of order 1 is referred to as a simple pole.

The point z_0 is called an essential isolated singularity of f if it is an isolated singularity of f but is neither a removable singularity nor a pole of f. Thus, z_0 is an essential singularity of f if and only if the principal part of the Laurent series of f has infinitely many nonvanishing terms (i.e., there are infinitely many negative integers n such that $a_n \neq 0$). An example is provided by the function $e^{1/z}$, defined in $\mathbf{C} \backslash \{0\}$. Its Laurent expansion about 0 is $\sum_{n=1}^{\infty} \frac{1}{n!} z^{-n}$. It thus has the origin as an essential singularity.

One says that the function f has an isolated singularity at ∞ if the open set G where f is defined contains a punctured disk centered at ∞ (that is, the exterior of some circle $|z| = R$, with $R \geq 0$). One then calls ∞ a removable singularity, a pole of order m, or an essential singularity of f if the function $f\left(\frac{1}{z}\right)$ has the origin as a removable singularity, a pole of order m, or an essential singularity, respectively. For example, the point ∞ is a removable singularity of the function $\frac{1}{z}$, a pole of order m of the function z^m, and an essential singularity of the function e^z.

Exercise VIII.7.1. Prove that if the holomorphic function f has an isolated singularity at z_0, then the principal part of the Laurent series of f at z_0 converges in $\mathbf{C} \backslash \{z_0\}$.

Exercise VIII.7.2. Locate and classify the isolated singularities of the following functions. (Don't ignore ∞ as a possible singularity.)

$$(a)\ \frac{z^5}{1+z+z^2+z^3+z^4}, \qquad (b)\ \frac{1}{\sin^2 z}, \qquad (c)\ \sin\frac{1}{z}.$$

Exercise VIII.7.3. Prove that a rational function has no essential singularities.

Exercise* VIII.7.4. (Partial fraction decomposition.) Let f be a nonconstant rational function, and let $z_1, z_2, \ldots, z_p$ be its poles in $\overline{\mathbf{C}}$. Prove that f can be written as $f = f_1 + f_2 + \cdots + f_p$, where each f_j is a rational function whose only pole is z_j.

Exercise VIII.7.5. Let the holomorphic function f be defined in $\mathbf{C}\backslash F$ where F is a finite set. Assume f has no essential singularities. Prove that f is a rational function.

Exercise VIII.7.6. Let the holomorphic functions f and g have poles of the same order, m, at the point z_0. Prove that

$$\lim_{z\to z_0}\frac{f(z)}{g(z)} = \lim_{z\to z_0}\frac{f'(z)}{g'(z)}$$

(a complex version of l'Hospital's rule).

Exercise VIII.7.7. The Bernoulli numbers B_n ($n = 0, 1, 2, \ldots$) are defined by

$$\frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n \qquad (|z| < 2\pi).$$

(As shown in this section, the function on the left has a removable singularity at the origin and so is represented by a power series about the origin. The radius of convergence of that series is 2π because the zeros of $e^z - 1$ closest to 0, other than 0 itself, are $\pm 2\pi i$.)

(a) Prove that $B_0 = 1$, and establish the recurrence relation

$$\sum_{k=0}^{n} \binom{n+1}{k} B_k = 0, \qquad n = 1, 2, \ldots,$$

which expresses B_n in terms of $B_0, \ldots, B_{n-1}$. Here, $\binom{n+1}{k}$ stands for the binomial coefficient $\dfrac{(n+1)!}{k!(n+1-k)!}$.

(b) Prove that $B_1 = -\dfrac{1}{2}$.

(c) Prove that $B_n = 0$ if n is odd and larger than 1.

(d) Prove that the numbers B_n are rational.

(e) Calculate $B_2,\ B_4,\ B_6,\ B_8,\ B_{10},\ B_{12}$.

VIII.8. Criterion for a Removable Singularity

Let the holomorphic function f have an isolated singularity at the point z_0 of $\mathbf{C}$. If f is bounded in some punctured disk with center z_0, then z_0 is a removable singularity of f.

In fact, take positive numbers M and ϵ such that f is defined and bounded in absolute value by M in the punctured disk $0 < |z - z_0| < \epsilon$. Let $\sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$ be the Laurent expansion of f in that punctured disk. For $0 < r < \epsilon$ let C_r denote the circle $|z - z_0| = r$, oriented counterclockwise. For $n > 0$ we have

$$a_{-n} = \frac{1}{2\pi i}\int_{C_r} (z - z_0)^{n-1} f(z)dz$$

(see Section VIII.2). The integrand in the preceding integral is bounded in absolute value by $r^{n-1}M$, so

$$|a_{-n}| \leq \frac{r^{n-1}M}{2\pi} L(C_r) = r^n M, \qquad n = 1, 2, \ldots.$$

Since r can be taken arbitrarily small it follows that $a_{-n} = 0$ for $n > 0$, the desired conclusion.

The preceding criterion obviously applies also in the case $z_0 = \infty$.

Exercise VIII.8.1. Show that Liouville's theorem is a corollary of the preceding criterion.

Exercise VIII.8.2. Let the holomorphic function f have an isolated singularity at the point z_0 of $\mathbf{C}$. Assume that there are positive numbers M and ϵ and a positive integer m such that $|f(z)| \leq M|z - z_0|^{-m}$ for $0 < |z - z_0| < \epsilon$. Prove that z_0 is either a removable singularity of f or a pole of order at most m.

Exercise VIII.8.3. Stand straight with feet about one meter apart, hands on hips. Bend at the waist, knees slightly flexed, and touch your left foot with right hand. Straighten. Bend again and touch right foot with left hand. Straighten. Repeat 15 times.

VIII.9. Criterion for a Pole

If the holomorphic function f has an isolated singularity at the point z_0, and if $\lim\limits_{z \to z_0} |f(z)| = \infty$, then z_0 is a pole of f.

In fact, under the given hypothesis, the point z_0 is a removable singularity of the function $g = \dfrac{1}{f}$, by VIII.8, and $g(z_0) = 0$. The desired conclusion follows immediately from this.

VIII.10. Casorati-Weierstrass Theorem

Let the holomorphic function f have an essential isolated singularity at the point z_0. Then, for any complex number w, there is a sequence $(z_n)_{n=1}^{\infty}$ in the region where f is defined such that $\lim_{n\to\infty} z_n = z_0$ and $\lim_{n\to\infty} f(z_n) = w$.

In fact, if the conclusion were to fail for a particular w, it would follow by VIII.8 that z_0 is a removable singularity of the function $g = \dfrac{1}{f-w}$. But that would mean z_0 is either a removable singularity of f (the case $g(z_0) \neq 0$) or a pole of f (the case $g(z_0) = 0$). The desired conclusion thus follows by contradiction.

VIII.11. Picard's Theorem

The Casorati-Weierstrass theorem reveals a wildness in the behavior of a holomorphic function near an essential isolated singularity: in any punctured disk centered at the singularity, the function comes arbitrarily close to every complex value. A deeper theorem of E. Picard reveals more: in any punctured disk centered at an essential isolated singularity, a holomorphic function actually assumes every complex value infinitely often, with perhaps one exception. In other words, except possibly for one value of w, the equality $\lim_{n\to\infty} f(z_n) = w$ in the conclusion of the Casorati-Weierstrass theorem can be replaced by $f(z_n) = w$, $n = 1, 2, \ldots$. Picard's theorem is illustrated by the function e^z, which is easily seen to assume every value other than 0 infinitely often in any punctured disk centered at ∞.

Picard's theorem, as it is considerably less elementary than the theorem of Casorati-Weierstrass, is ordinarily not proved in introductory courses. The reader can find proofs in the books of Ahlfors, Burckel, Conway, and Saks and Zygmund.

VIII.12. Residues

Let the holomorphic function f have an isolated singularity at the point z_0 of $\mathbf{C}$. The residue of f at z_0, denoted by $\operatorname{res}_{z_0} f$, or by $\operatorname{res}_{z=z_0} f(z)$, is defined to be the coefficient of $(z-z_0)^{-1}$ in the Laurent expansion of f near z_0. The residues of a holomorphic function can be used in the evaluation of integrals of the function. This is the content of the residue theorem, which will be given in Chapter X. The simplest case appears already in Section VIII.2: if C is a counterclockwise oriented circle centered at z_0 and contained with its punctured interior in the region where f is defined, then

$$\int_C f(z)dz \;=\; 2\pi i \operatorname{res}_{z_0} f.$$

Here are a few examples.

Example 1. The Laurent expansion of $e^{1/z}$ about 0 is $\sum_{n=0}^{\infty} \frac{1}{n!} z^{-n}$, giving $\operatorname{res}_{z=0} e^{1/z} = 1$.

Example 2. $f(z) = \dfrac{z^3}{z-1}, \quad z_0 = 1.$

The Taylor expansion of z^3 about the point 1 is

$$z^3 = 1 + 3(z-1) + 3(z-1)^2 + (z-1)^3.$$

Hence, the Laurent expansion of $\dfrac{z^3}{z-1}$ about the point 1 is

$$\frac{z^3}{z-1} = \frac{1}{z-1} + 3 + 3(z-1) + (z-1)^2,$$

giving

$$\operatorname{res}_{z=1} \frac{z^3}{z-1} = 1.$$

Example 3. $f(z) = \dfrac{z^3}{(z-1)^2}, \quad z_0 = 1.$

The same calculation as used in the last example shows that the Laurent expansion of $\dfrac{z^3}{(z-1)^2}$ about the point 1 is

$$\frac{z^3}{(z-1)^2} = \frac{1}{(z-1)^2} + \frac{3}{z-1} + 3 + (z-1),$$

giving

$$\operatorname{res}_{z=1} \frac{z^3}{(z-1)^2} = 3.$$

Exercise* VIII.12.1. Let the functions g and h be holomorphic in an open set containing the point z_0, and assume that h has a simple zero at z_0. Prove that

$$\operatorname{res}_{z_0} \frac{g}{h} = \frac{g(z_0)}{h'(z_0)}.$$

Exercise VIII.12.2. Determine the residues of the following functions at each of their isolated singularities in **C**.

(a) $\dfrac{z^p}{1-z^q}$ (p, q positive integers),

(b) $\dfrac{z^5}{(z^2-1)^2}$,

(c) $\dfrac{\cos z}{1+z+z^2}$,

(d) $\dfrac{1}{\sin z}$.

Exercise* VIII.12.3. Let the function f be holomorphic in an open set containing the point z_0 and have a zero at z_0 of order m. Prove that

$$\operatorname{res}_{z_0} \frac{f'}{f} = m.$$

Chapter IX

Cauchy's Theorem

A general version of Cauchy's theorem will be established in Section IX.10. The necessary analytic machinery has already been developed. Still needed are some topological preliminaries centering on the notion of winding number. A second version of Cauchy's theorem, involving homotopy, is given in Section IX.14. This will entail some additional topological preliminaries.

At the end of the chapter an approximation theorem, Runge's theorem, is established and used to give an alternative proof of Cauchy's theorem. This material is not needed in the final chapter, Chapter X.

IX.1. Continuous Logarithms

Let ϕ be a continuous, nowhere vanishing, complex-valued function on the subinterval $[a, b]$ of $\mathbf{R}$. Then there is a continuous function ψ on $[a, b]$ such that $\phi = e^{\psi}$. The function ψ is unique to within addition of a constant integer multiple of $2\pi i$.

As for uniqueness, if ψ_1 and ψ_2 are two functions with the required properties, then the function $\frac{1}{2\pi i}(\psi_1 - \psi_2)$ is continuous and integer valued, hence constant (since $[a, b]$ is connected).

We establish the existence of ψ first for the case where the range of ϕ lies in an open half-plane H bounded by a line through the origin. In that case, as we know from Section IV.10, there is a branch l of $\log z$ in H. The composite function $\psi = l \circ \phi$ then has the required properties. Note that, by suitably choosing the branch l of $\log z$, we can arrange to make $\psi(a)$ equal to any preassigned value of $\log \phi(a)$.

To establish the existence of ψ in the general case, we partition $[a,b]$ into subintervals $[t_0,t_1]$, $[t_1,t_2],\ldots,[t_{p-1},t_p]$, where $a=t_0<t_1<\cdots<t_p=b$, such that the range of ϕ on each subinterval lies in an open half-plane bounded by a line through the origin. This is easily done on the basis of the continuity of ϕ and the compactness of $[a,b]$. By the special case already treated, there is a continuous function ψ_1 on $[t_0,t_1]$ such that $\phi=e^{\psi_1}$ on that interval. By the same special case, there is a continuous function ψ_2 on $[t_1,t_2]$ such that $\phi=e^{\psi_2}$ on that interval, and such that $\psi_2(t_1)=\psi_1(t_1)$. Continuing in this way, we obtain on each interval $[t_{j-1},t_j]$ a continuous function ψ_j such that $\phi=e^{\psi_j}$ there, and $\psi_j(t_{j-1})=\psi_{j-1}(t_{j-1})$. The function ψ on $[a,b]$ that equals ψ_j on $[t_{j-1},t_j]$ $(j=1,\ldots,p)$ then has the required properties.

IX.2. Piecewise C^1 Case

Assume that the function ϕ *in* IX.1 *is piecewise* C^1. *Let* c *be any value of* $\log\phi(a)$. *Then the function* ψ *defined by*

$$\psi(t) \;=\; c\;+\;\int_a^t \frac{\phi'(s)}{\phi(s)}ds, \qquad a\le t\le b,$$

has the required properties: it is continuous and $\phi=e^{\psi}$.

The argument is standard. First, we may as well assume ϕ is of class C^1, because the piecewise-C^1 case is an easy consequence of that one. If ϕ is of class C^1, then the function ψ defined above is differentiable, with $\psi'=\dfrac{\phi'}{\phi}$. Then

$$\left(\phi e^{-\psi}\right)' \;=\; \phi' e^{-\psi}-\phi\psi' e^{-\psi}=0,$$

implying that the function $\phi e^{-\psi}$ is constant. Since $\phi(a)=e^{\psi(a)}$, the constant is 1, and the proof is complete.

IX.3. Increments in the Logarithm and Argument Along a Curve

Let $\gamma:[a,b]\to\mathbf{C}$ be a curve in $\mathbf{C}$, and let f be a continuous complex-valued function that is defined and nowhere vanishing on γ. By IX.1 there is a continuous function ψ on $[a,b]$, unique to within addition of a constant integer multiple of $2\pi i$, such that $f\circ\gamma=e^{\psi}$. The increment in $\log f$ on γ, denoted $\Delta(\log f,\gamma)$, is defined to be the difference $\psi(b)-\psi(a)$. The increment in $\arg f$ on γ, denoted $\Delta(\arg f,\gamma)$, is defined to be $\operatorname{Im}\Delta(\log f,\gamma)$. In the special case where γ is a closed curve, the real part of $\Delta(\log f,\gamma)$ vanishes (since all logarithms of the same number have the same real part), so $\Delta(\arg f, \gamma) = \dfrac{1}{i}\Delta(\log f, \gamma)$. Moreover, if γ is closed then $\Delta(\arg f, \gamma)$ is an integer multiple of 2π.

The intuition is as follows. One can imagine a perceptive point moving along γ from $\gamma(a)$ to $\gamma(b)$, and, as it proceeds, keeping track in a continuous way of a value of $\log f(z)$. The difference between the terminal and initial values perceived by the point is the increment in $\log f$ along γ.

If the curve γ is piecewise C^1 and the function f is holomorphic in an open set containing γ (and nonvanishing on γ), then by IX.2 we have

$$\Delta(\log f, \gamma) = \int_a^b \frac{f'(\gamma(t))\gamma'(t)}{f(\gamma(t))} dt,$$

which can be rewritten

$$\Delta(\log f, \gamma) = \int_\gamma \frac{f'(z)}{f(z)} dz.$$

If in addition γ is a closed curve, then the real part of $\Delta(\log f, \gamma)$ is 0, and we have

$$\Delta(\arg f, \gamma) = \frac{1}{i}\int_\gamma \frac{f'(z)}{f(z)} dz.$$

IX.4. Winding Number

Let $\gamma : [a, b] \to \mathbf{C}$ be a closed curve in $\mathbf{C}$, and let z_0 be a point in the complement of γ. The function $f(z) = z - z_0$ is then nonvanishing on γ, so the increment $\Delta(\arg(z - z_0), \gamma)$ is defined by the discussion in the preceding section. That increment is an integer multiple of 2π. The integer $\dfrac{1}{2\pi}\Delta(\arg(z - z_0), \gamma)$ is called the winding number of γ about z_0, or the index of z_0 with respect to γ, and denoted by $\operatorname{ind}_\gamma(z_0)$. If one imagines an observer stationed at z_0 watching a point move once around γ, and continuously turning so as always to face the point, then $\operatorname{ind}_\gamma(z_0)$ will equal the net number of revolutions performed by the observer, counterclockwise revolutions being reckoned as positive and clockwise revolutions as negative. The winding number of a curve about a point, at least in simple cases, can be read off from a picture of the curve. As we shall see shortly, it takes a constant value on each connected component of the complement of the curve. In the examples in Figure 6, the winding number is indicated for each complementary component.

IX.5. Case of a Piecewise-C^1 Curve

Let $\gamma : [a, b] \to \mathbf{C}$ be a piecewise-C^1 closed curve in $\mathbf{C}$, and let z_0 be a point in the complement of γ. Applying the formula at the end of Section IX.3 to

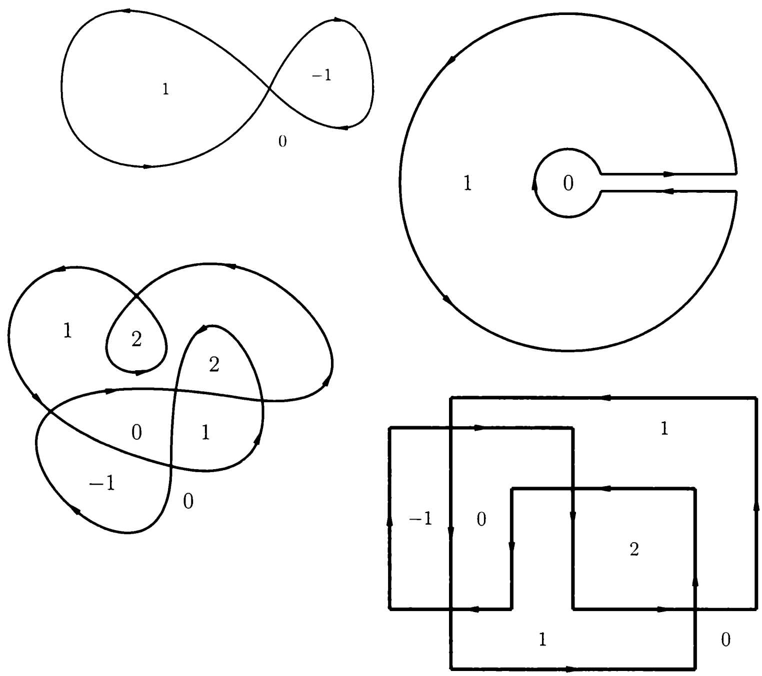

Figure 6. Winding Numbers

the function $f(z) = z - z_0$, we obtain an integral expression for the winding number of γ about z_0:

$$\operatorname{ind}_\gamma(z_0) = \frac{1}{2\pi i}\int_\gamma \frac{1}{z - z_0}dz.$$

The quantity on the right side of the preceding equality is easily seen to be a continuous function of z_0 on $\mathbf{C}\backslash\gamma$. (In fact, by what is proved in Section VII.7, it is a holomorphic function of z_0.) Since it is also integer valued, it must be constant on each connected component of $\mathbf{C}\backslash\gamma$. Because γ is bounded, only one of the components of $\mathbf{C}\backslash\gamma$ is unbounded. As $z_0 \to \infty$, the integral in the expression above for $\operatorname{ind}_\gamma(z_0)$ tends to 0 (by standard estimates; see Section VI.10), so also $\operatorname{ind}_\gamma(z_0) \to 0$. Since $\operatorname{ind}_\gamma(z_0)$ is an integer, it must actually equal 0 when z_0 is sufficiently large. Since it takes a constant value on each component of $\mathbf{C}\backslash\gamma$, it therefore equals 0 on the unbounded component of $\mathbf{C}\backslash\gamma$.

Exercise IX.5.1. Prove that the properties of winding number established in the preceding paragraph for piecewise-C^1 curves hold for general curves: if $\gamma : [a, b] \to \mathbf{C}$ is a closed curve in $\mathbf{C}$, then $\operatorname{ind}_\gamma(z_0)$ is constant on each component of $\mathbf{C}\backslash\gamma$ and 0 on the unbounded component.

Exercise IX.5.2. Prove the residue theorem for rational functions: Let f be a rational function, and let $z_1, \dots, z_p$ be the poles of f in $\mathbf{C}$. Let γ be a piecewise-C^1 closed curve that does not pass through any of the poles of f. Then

$$\frac{1}{2\pi i}\int_\gamma f(z)dz = \sum_{j=1}^{p} \operatorname{ind}_\gamma(z_j)\cdot \operatorname{res}_{z_j} f.$$

Exercise IX.5.3. Prove that if z_0 is a point in $\mathbf{C}\backslash\mathbf{R}$, then

$$\lim_{R\to\infty}\frac{1}{2\pi i}\int_{-R}^{R}\frac{1}{t-z_0}dt$$

is $\frac{1}{2}$ if $\operatorname{Im} z_0 > 0$ and $-\frac{1}{2}$ if $\operatorname{Im} z_0 < 0$. (Intuitively, the real axis winds $\frac{1}{2}$ time around each point in the upper half-plane and $-\frac{1}{2}$ time around each point in the lower half-plane.)

IX.6. Contours

The term "contour" is commonly used in an informal way in complex function theory to refer to a closed integration path. Here we shall assign the term a more specific meaning. By a contour we shall mean a formal sum $\Gamma = \sum_{j=1}^{p} n_j\gamma_j$, where $\gamma_1, \dots, \gamma_p$ are piecewise C^1 closed curves in $\mathbf{C}$ and $n_1, \dots, n_p$ are integers. If f is a continuous complex-valued function defined on each γ_j—or, as we shall say, defined on Γ—we define the complex integral of f over Γ by

$$\int_\Gamma f(z)dz = \sum_{j=1}^{p} n_j\int_{\gamma_j} f(z)dz.$$

We shall identify a piecewise C^1 closed curve γ with the contour 1γ. In general, we shall identify two contours if they are indistinguishable for purposes of integration. Thus, for example, two piecewise C^1 closed curves will be identified if one of them is a reparametrization of the other (something we have been doing all along). If γ is a piecewise C^1 closed curve, the contour -1γ will be identified with $-\gamma$, the reverse of γ (see Section (VI.9)). If γ is defined on the parameter interval $[a, b]$, then the contour 2γ can be identified with the curve on the parameter interval $[a, 2b-a]$ that takes the value $\gamma(t)$ at a point t in $[a, b]$ and the value $\gamma(t-b+a)$ at a point t in $[b, 2b-a]$.

If $\Gamma = \sum_{j=1}^{p} n_j\gamma_j$ and $\Gamma' = \sum_{j=1}^{p} n_j'\gamma_j$ are two contours, we define their sum by

$$\Gamma + \Gamma' = \sum_{j=1}^{p}(n_j + n_j')\gamma_j.$$

Then, if f is a continuous function defined on every γ_j, we have

$$\int_{\Gamma+\Gamma'} f(z)dz = \int_{\Gamma} f(z)dz + \int_{\Gamma'} f(z)dz.$$

The algebraic structure associated with contours will not be formalized here. With sufficient effort one can define an equivalence relation on the family of contours and make the set of equivalence classes into an abelian group. To carry out this program here would, however, be a digression with little bearing on our central purpose.

IX.7. Winding Numbers of Contours

Let $\Gamma = \sum_{j=1}^{p} n_j\gamma_j$ be a contour, and assume that each n_j is nonzero. By the complement of Γ, written $\mathbf{C}\backslash\Gamma$, we shall mean the set of points of $\mathbf{C}$ that are in the complement of each γ_j. (As with curves, we shall, when convenient, speak of contours as if they were subsets of $\mathbf{C}$. This will cause no confusion in practice.) If z_0 is such a point, we define the winding number of Γ about z_0, or the index of z_0 with respect to Γ, by

$$\mathrm{ind}_\Gamma(z_0) = \sum_{j=1}^{p} n_j \ \mathrm{ind}_{\gamma_j}(z_0).$$

As in the case of a single closed curve, we have the integral formula

$$\mathrm{ind}_\Gamma(z_0) = \frac{1}{2\pi i}\int_\Gamma \frac{1}{z-z_0}dz.$$

The reasoning used in Section IX.5 for single curves applies without change to contours; it shows that $\mathrm{ind}_\Gamma(z_0)$ is constant on each connected component of $\mathbf{C}\backslash\Gamma$ and equals 0 on the unbounded component.

As a simple example, let C_1 be the circle $|z|=1$ and C_2 the circle $|z|=2$, both oriented counterclockwise. For the contour $\Gamma = C_2 - C_1$ we have

$$\mathrm{ind}_\Gamma(z_0) = \begin{cases} 1 & \text{for } 1<|z_0|<2 \\ 0 & \text{for } |z_0|<1 \text{ and for } |z_0|>2. \end{cases}$$

We define the interior of the contour Γ, denoted $\operatorname{int}\Gamma$, to be the set of points z_0 in $\mathbf{C}\backslash\Gamma$ such that $\mathrm{ind}_\Gamma(z_0)\neq 0$, and the exterior, denoted $\operatorname{ext}\Gamma$, to be the set of points z_0 in $\mathbf{C}\backslash\Gamma$ such that $\mathrm{ind}_\Gamma(z_0)=0$. Each of these sets is a union of certain of the connected components of $\mathbf{C}\backslash\Gamma$. As those components are open sets, the sets $\operatorname{int}\Gamma$ and $\operatorname{ext}\Gamma$ are open. Because the unbounded component of $\mathbf{C}\backslash\Gamma$ is contained in $\operatorname{ext}\Gamma$, the set $\operatorname{int}\Gamma$ is bounded. All boundary points of $\operatorname{int}\Gamma$ and of $\operatorname{ext}\Gamma$ are contained in Γ.

A contour Γ will be called simple if $\mathrm{ind}_\Gamma(z_0)$ is either 0 or 1 for each point z_0 in $\mathbf{C}\backslash\Gamma$.

IX.8. Separation Lemma

Let G be an open subset of $\mathbf{C}$ *and K a compact subset of G. Then there is a simple contour Γ in $G \backslash K$ such that $K \subset \operatorname{int}\Gamma \subset G$.*

In other words, K can be separated from the complement of G by a simple contour. This is our main topological preliminary to Cauchy's theorem. We shall prove it by forming a grid in the plane and building the required contour out of selected segments of the grid.

Let δ be a positive number less than $\operatorname{dist}(K, \mathbf{C}\backslash G)$ $(= \inf\{|z-w| : z \in K,\ w \in \mathbf{C}\backslash G\})$. By means of equally spaced horizontal and vertical lines, we subdivide $\mathbf{C}$ into nonoverlapping squares each of diameter δ. By a square here we mean a solid square, consisting of both boundary and interior. The term "nonoverlapping" expresses the fact that distinct squares in the subdivision have disjoint interiors. (Two distinct squares, if not disjoint, share at most a common edge or a common vertex.)

Let $S_1, S_2, \ldots, S_q$ be an enumeration of those squares in the subdivision that intersect K. The collection $\{S_1, S_2, \ldots, S_q\}$ will be denoted by $\mathcal{S}$. By the choice of δ, each square in $\mathcal{S}$ is contained in G. For $j = 1, \ldots, q$ we let σ_j denote the counterclockwise oriented boundary of the square S_j.

Each square in our subdivision has four edges. We let $\mathcal{E}$ be the collection of those edges that belong to only one square in the collection $\mathcal{S}$. Each edge in $\mathcal{E}$ is then contained in G but disjoint from K. We give each edge in $\mathcal{E}$ the orientation that it inherits from the boundary of the square in $\mathcal{S}$ that contains it. We shall refer to an edge that belongs to two squares in $\mathcal{S}$ as an interior edge.

We shall show first that the edges in $\mathcal{E}$ can be organized into a contour, in other words, that there is a contour Γ such that

$$\sum_{\mathrm{E}\in\mathcal{E}} \int_{\mathrm{E}} f(z)dz = \int_{\Gamma} f(z)dz$$

for any function f that is continuous on each edge in $\mathcal{E}$. Subsequently we shall show that Γ possesses the other required properties.

A sequence $(\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_p)$ of edges in $\mathcal{E}$, without repetitions, will be called a chain if the terminal point of E_{k-1} coincides with the initial point of E_k $(k = 2, \ldots, p)$. Such a chain will be called a cycle if, in addition, the terminal point of E_p coincides with the initial point of E_1. If $(\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_p)$ is a cycle, we can build from it a piecewise-C^1 closed curve γ in an obvious way: for $k = 1, \ldots, p$ we parametrize E_k (linearly, say) by means of a function γ_k on the interval $[k-1, k]$, and then we let γ be the function on $[0, p]$

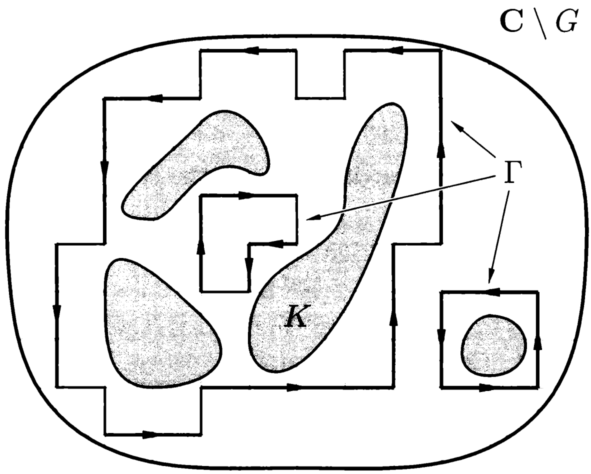

Figure 7. Construction of Γ in the proof of the separation lemma. The set K in this example has four components.

that coincides with γ_k on $[k-1, k]$ for each k. With γ so defined, we have

$$\sum_{k=1}^{p} \int_{\mathrm{E}_k} f(z)dz = \int_{\gamma} f(z)dz$$

for any continuous function f defined on each E_k. To construct the desired contour Γ, therefore, it will suffice to show that we can decompose $\mathcal{E}$ into a disjoint union of cycles.

To accomplish the latter we note the following property of $\mathcal{E}$, which one can verify by examining all possible cases:

($*$) Each vertex in our subdivision is the initial point of the same number of edges in $\mathcal{E}$ as it is the terminal point of. (The number is either 0, 1, or 2.)

From ($*$) we can conclude that a maximal chain in $\mathcal{E}$, that is, a chain we cannot lengthen by appending an edge from $\mathcal{E}$ at its beginning or end, must in fact be a cycle. Since $\mathcal{E}$, being finite, obviously contains maximal chains, it contains cycles. If we remove a cycle from $\mathcal{E}$ we obtain a smaller collection of edges which will still satisfy ($*$). If the smaller collection is nonempty we can extract a cycle from it, and so on. After finitely many steps we obtain in this way the desired decomposition of $\mathcal{E}$ into a disjoint union of cycles, completing the construction of the contour Γ.

It remains to show that Γ has the required properties. That Γ is contained in $G \backslash K$ is obvious. We need to show that Γ is simple and that $K \subset \operatorname{int} \Gamma \subset G$.

Suppose z_0 is a point in the interior of one of the squares in $\mathcal{S}$, say the square S_{j_0}. Then

$$\frac{1}{2\pi i}\int_{\sigma_j}\frac{1}{z-z_0}dz = \begin{cases}1, & j = j_0\\ 0, & j \neq j_0.\end{cases}$$

It follows that

$$\frac{1}{2\pi i}\sum_{j=1}^{q}\int_{\sigma_j}\frac{1}{z-z_0}dz \;=\; 1.$$

The left side in the preceding equality equals $\frac{1}{2\pi i}\int_{\Gamma}\frac{1}{z-z_0}dz$. To see this, note that in the sum of integrals that appears, each of the interior edges (that is, the edges belonging to two of the squares in $\mathcal{S}$) makes two contributions which cancel (since the orientations such an edge inherits from the two squares containing it are opposite), leaving only the contributions from the edges in $\mathcal{E}$. We thus have

$$\frac{1}{2\pi i}\int_{\Gamma}\frac{1}{z-z_0}dz \;=\; 1$$

whenever z_0 is a point in the interior of one of the squares in $\mathcal{S}$. The expression on the left side is a continuous function of z_0 on the complement of Γ, so the equality continues to hold if z_0 is a point off Γ that lies on one of the interior edges. In particular, we have $\mathrm{ind}_{\Gamma}(z_0) = 1$ for all z_0 in K.

On the other hand, if the point z_0 is not in any of the squares in $\mathcal{S}$, then $\int_{\sigma_j}\frac{1}{z-z_0}dz = 0$ for every j, and, reasoning as above, we can conclude that $\mathrm{ind}_{\Gamma}(z_0) = 0$. This applies, in particular, for z_0 in $\mathbf{C}\backslash G$.

We have thus shown that $\mathrm{ind}_{\Gamma}(z_0)$ takes only the values 0 and 1, that it is 1 for z_0 in K, and that it is 0 for z_0 in $\mathbf{C}\backslash G$. The proof of the lemma is complete.

IX.9. Addendum to the Separation Lemma

Let Γ be the contour constructed above, and let f be a holomorphic function in G. Then

$$f(z_0) \;=\; \frac{1}{2\pi i}\int_{\Gamma}\frac{f(z)}{z-z_0}dz, \qquad z_0 \in K.$$

This special case of Cauchy's formula will be used to deduce the general version of Cauchy's theorem in the next section. To prove it we note first that $\int_{\sigma_j} f(z)dz = 0$ for each j, by Cauchy's theorem for a square (an immediate corollary of Cauchy's theorem for a triangle). Since $\sum_{j=1}^{q}\int_{\sigma_j} f(z)dz =$

$\int_\Gamma f(z)dz$ (by the same reasoning as used in the proof of the separation lemma), we have $\int_\Gamma f(z)dz = 0$.

Now fix a point z_0 in K, and define the function g in G by

$$g(z) = \begin{cases} \dfrac{f(z) - f(z_0)}{z - z_0}, & z \neq z_0 \\ f'(z_0), & z = z_0. \end{cases}$$

Then g is holomorphic, so, by what was just proved, $\int_\Gamma g(z)dz = 0$. Consequently

$$\int_\Gamma \frac{f(z)}{z - z_0} dz = f(z_0) \int_\Gamma \frac{1}{z - z_0} dz = 2\pi i f(z_0),$$

as desired.

IX.10. Cauchy's Theorem

Let G be an open subset of $\mathbf{C}$, and let Γ be a contour contained with its interior in G. Then $\int_\Gamma f(z)dz = 0$ for every function f that is holomorphic in G.

In order to prove this, let K be the complement in $\mathbf{C}$ of the exterior of Γ. Because ext Γ is an open set, the set K is closed. Also $\mathbf{C}\backslash K$ contains $\mathbf{C}\backslash G$ as well as the unbounded component of $\mathbf{C}\backslash\Gamma$, whose complement is bounded (by the boundedness of Γ). It follows that K is a compact subset of G. Obviously $\Gamma \subset K$.

By the separation lemma and its addendum, there is a simple contour Γ_1 in $G\backslash K$ such that $K \subset \operatorname{int}\Gamma_1 \subset G$, and such that, for any function f that is holomorphic in G,

$$f(z) = \frac{1}{2\pi i} \int_{\Gamma_1} \frac{f(\zeta)}{\zeta - z} d\zeta, \qquad z \in K.$$

For such a function f we thus have

$$\int_\Gamma f(z)dz = \frac{1}{2\pi i} \int_\Gamma \int_{\Gamma_1} \frac{f(\zeta)}{\zeta - z} d\zeta\, dz.$$

In the double integral on the right, it is permissible to change the order of integration because the integrand is a continuous function of the pair of variables z, ζ. (More precisely, if one writes out the double contour integral in terms of real parametrizing variables for the constituent curves of Γ and Γ_1, one finds that the integrands that occur are continuous functions of the parameters on the appropriate rectangles in $\mathbf{R}^2$.) Hence

$$\begin{aligned}\int_\Gamma f(z)dz &= \frac{1}{2\pi i}\int_{\Gamma_1}\left[f(\zeta)\int_\Gamma \frac{1}{\zeta - z}dz\right]d\zeta \\ &= -\int_{\Gamma_1} f(\zeta)\,\mathrm{ind}_\Gamma(\zeta)d\zeta.\end{aligned}$$

The right side is 0 because Γ_1 is contained in the complement of K, so that $\mathrm{ind}_\Gamma(\zeta) = 0$ for all ζ on Γ_1. This completes the proof of Cauchy's theorem.

Exercise IX.10.1. Justify in detail the change in order of integration in the preceding proof.

Exercise IX.10.2. Explain why VII.2, Cauchy's theorem for a convex region, is a special case of IX.10, the general Cauchy theorem. (It is not a bona fide corollary because it was used in the developments leading to IX.10.)

IX.11. Homotopy

Let G be an open subset of $\mathbf{C}$. Two closed curves $\gamma_0 : [0,1] \to G$ and $\gamma_1 : [0,1] \to G$ are said to be homotopic in G if there is a continuous map $\gamma : [0,1] \times [0,1] \to G$ satisfying

(i) $\gamma(t,0) = \gamma_0(t)$ for all t,

(ii) $\gamma(t,1) = \gamma_1(t)$ for all t,

(iii) $\gamma(0,s) = \gamma(1,s)$ for all s.

Intuitively, the curves γ_s, $0 \le s \le 1$, defined by $\gamma_s(t) = \gamma(t,s)$, provide a continuous deformation within G from γ_0 to γ_1. It is clear that γ_0 and γ_1 can only be homotopic in G if they lie in the same connected component of G.

It will be shown that if γ_0 and γ_1 are homotopic in G then they have the same winding number about every point in $\mathbf{C}\backslash G$. This will depend on a two-dimensional version of the result on continuous logarithms in Section IX.1.

IX.12. Continuous Logarithms—2-D Version

Let $\phi : [0,1] \times [0,1] \to \mathbf{C}\backslash\{0\}$ be a continuous function. Then there is a continuous function $\psi : [0,1] \times [0,1] \to \mathbf{C}$ such that $\phi = e^\psi$. The function ψ is unique to within addition of an integer multiple of $2\pi i$.

This will be deduced from IX.1, its one-dimensional counterpart. The statement about uniqueness follows in the same way as did the corresponding statement in IX.1.

Define $\phi_{00} : [0,1] \to \mathbf{C}\backslash\{0\}$ by $\phi_{00}(t) = \phi(t,0)$. For each t in $[0,1]$, define $\phi_t : [0,1] \to \mathbf{C}\backslash\{0\}$ by $\phi_t(s) = \phi(t,s)$. By IX.1, there is a continuous function $\psi_{00} : [0,1] \to \mathbf{C}$ such that $\phi_{00} = e^{\psi_{00}}$. By IX.1, again, there is for each t in $[0,1]$ a continuous function $\psi_t : [0,1] \to \mathbf{C}$ such that $\phi_t = e^{\psi_t}$ and $\psi_t(0) = \psi_{00}(t)$. Define $\psi : [0,1] \times [0,1] \to \mathbf{C}$ by $\psi(t,s) = \psi_t(s)$. The equality $\phi = e^{\psi}$ is immediate. It remains to prove that ψ is continuous.

The range of ϕ is a compact subset of $\mathbf{C}\backslash\{0\}$ and so has a positive distance from 0. Choose $\epsilon > 0$ such that $|\phi(t,s)| > \epsilon$ for all (t,s). Fix t_0 in $[0,1]$. By the uniform continuity of ϕ, there is a $\delta > 0$ such that $|\phi(t,s) - \phi(t_0,s)| < \epsilon$ for all s whenever $|t - t_0| \leq \delta$. Thus, for $|t - t_0| \leq \delta$ we have

$$\left|\frac{\phi(t,s)}{\phi(t_0,s)} - 1\right| = \frac{1}{\phi(t_0,s)}|\phi(t,s) - \phi(t_0,s)| < \frac{|\phi(t,s) - \phi(t_0,s)|}{\epsilon} < 1,$$

implying that

$$\operatorname{Re}\frac{\phi(t,s)}{\phi(t_0,s)} > 0.$$

Define the function χ on the rectangle

$$R = \big([t_0 - \delta, t_0 + \delta] \cap [0,1]\big) \times [0,1]$$

by $\chi(t,s) = \operatorname{Log}\frac{\phi(t,s)}{\phi(t_0,s)}$. Then χ is continuous, $\phi(t,s)/\phi(t_0,s) = e^{\chi(t,s)}$ on R, and $\chi(t_0,s) = 1$ for all s. The function $\widetilde{\psi}$ on R defined by $\widetilde{\psi}(t,s) = \chi(t,s) + \phi(t_0,s)$ is continuous and satisfies $\phi = \epsilon^{\widetilde{\psi}}$ on R. Also $\widetilde{\psi}(t_0,0) = \psi(t_0,0)$. On the interval $[t_0 - \delta, t_0 + \delta] \cap [0,1]$, the functions $t \mapsto \psi(t,0)$ and $t \mapsto \widetilde{\psi}(t,0)$ are continuous logarithms of the same function and they agree for $t = t_0$, so they coincide. For each t in $[t_0 - \delta, t_0 + \delta] \cap [0,1]$, then, the functions $s \mapsto \psi(t,s)$ and $s \mapsto \widetilde{\psi}(t,s)$ on $[0,1]$ are continuous logarithms of the same function that agree at $s = 0$, so they coincide. We see therefore that ψ coincides with the continuous function $\widetilde{\psi}$ on R; in particular, ψ is continuous at each point of $\{t_0\} \times [0,1]$. Our proof is complete.

IX.13. Homotopy and Winding Numbers

Let G be an open subset of $\mathbf{C}$. Let the closed curves γ_0 and γ_1 in G be homotopic in G. Then $\operatorname{ind}_{\gamma_0}(z) = \operatorname{ind}_{\gamma_1}(z)$ for all z in $\mathbf{C}\backslash G$.

To establish this, take $\gamma : [0,1] \times [0,1] \to G$ as in the definition in IX.11, and fix a point z_0 in $\mathbf{C}\backslash G$. Define the function $\phi : [0,1] \times [0,1] \to \mathbf{C}\backslash\{0\}$ by $\phi(t,s) = \gamma(t,s) - z_0$. By IX.12 there is a continuous function $\psi : [0,1] \times [0,1] \to \mathbf{C}$ such that $\phi = e^{\psi}$. Then for each s in $[0,1]$ we have

$$\operatorname{ind}_{\gamma_s}(z_0) = \frac{1}{2\pi i}\big(\psi(1,s) - \psi(0,s)\big),$$

showing that the function $s \mapsto \mathrm{ind}_{\gamma_s}(z_0)$ is continuous, hence constant. In particular, $\mathrm{ind}_{\gamma_0}(z_0) = \mathrm{ind}_{\gamma_1}(z_0)$, as desired.

IX.14. Homotopy Version of Cauchy's Theorem

Let the function f be holomorphic in the open set G, and let γ_0 and γ_1 be two piecewise-C^1 closed curves in G that are homotopic in G. Then $\int_{\gamma_0} f(z)dz = \int_{\gamma_1} f(z)dz$.

This follows from IX.10 because, by IX.13, the contour $\gamma_1 - \gamma_0$ has winding number 0 about all points of $\mathbf{C}\backslash G$.

IX.15. Runge's Approximation Theorem

Let K be a compact subset of $\mathbf{C}$, and let f be a holomorphic function defined in an open set containing K. Then f can be uniformly approximated on K by rational functions whose poles lie in $\mathbf{C}\backslash K$.

The last statement says that, for any positive number ϵ, there is a rational function g whose poles lie in $\mathbf{C}\backslash K$ such that $|f(z) - g(z)| < \epsilon$ for all z in K. An equivalent statement is that there exists a sequence of such rational functions that converges uniformly on K to f. Runge's theorem was published in 1885 (as was the more famous approximation theorem of K. Weierstrass).

To prove Runge's theorem we first note that the family of rational functions whose poles lie in any preassigned set is closed under addition: if two functions belong to the family, then so does their sum. It follows that the family of functions on K that can be uniformly approximated (on K) by such rational functions is also closed under addition.

Let the function f be holomorphic in the open set G containing K. Let Γ be the simple contour in $G\backslash K$ constructed in the proof of the separation lemma. By the addendum to the separation lemma,

$$f(z) = \frac{1}{2\pi i}\int_\Gamma \frac{f(\zeta)}{\zeta - z}d\zeta, \qquad z \in K.$$

We created Γ by piecing together certain directed edges produced by a subdivision of the plane into squares. Let $\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_s$ be an enumeration of those edges. We then have $f = \sum_{j=1}^s f_j$, where

$$f_j(z) = \frac{1}{2\pi i}\int_{\mathrm{E}_j} \frac{f(\zeta)}{\zeta - z}d\zeta, \qquad z \in K.$$

By the remark in the preceding paragraph, it will suffice to show that each f_j can be uniformly approximated on K by rational functions with poles in $\mathbf{C}\backslash K$.

Fix an index j. Let the segment E_j be partitioned into subsegments each having a length less than the distance between E_j and K. There is a corresponding expression for f_j as a sum of functions, each summand being $\frac{1}{2\pi i}$ times the Cauchy integral of f over one of the subsegments. By the same reasoning as was just used, it will suffice to show that each summand in that expression can be uniformly approximated on K by rational functions with poles in $\mathbf{C}\backslash K$. A typical summand has the form

$$h(z) = \frac{1}{2\pi i}\int_{\mathrm{E}} \frac{f(\zeta)}{\zeta - z} d\zeta,$$

where E is the subsegment in question. Thus, the length of E is less than the distance between E and K. Denote the preceding length by δ, and let z_0 be the midpoint of E. Then, as proved in Section VIII.3, the function h is represented in the region $|z - z_0| > \delta/2$ by the Laurent series $\sum_{n=-\infty}^{-1} a_n(z - z_0)^n$, where $a_n = \frac{-1}{2\pi i}\int_{\mathrm{E}} (\zeta - z_0)^{-n-1} f(\zeta) d\zeta$. The series converges locally uniformly in the preceding region, which contains the set K. The series therefore converges uniformly on K. As the partial sums of the series are rational functions whose only pole is z_0, this completes the proof of Runge's theorem.

IX.16. Second Proof of Cauchy's Theorem

With Runge's theorem one can give an alternative proof of Cauchy's theorem. Let G, Γ, and f be as in Section IX.11, namely, G is an open subset of $\mathbf{C}$, Γ is a contour contained with its interior in G, and f is a holomorphic function in G. As in the earlier proof, let K be the complement in $\mathbf{C}$ of the exterior of Γ. Then K is a compact subset of G, and $\Gamma \subset K$.

We first note that $\int_\Gamma g(z)dz = 0$ whenever g is a rational function whose poles lie in $\mathbf{C}\backslash K$. In fact, any such g is, to within addition of a constant, a linear combination of functions of the form $(z - z_0)^{-n}$, where z_0 is in $\mathbf{C}\backslash K$ and n is a positive integer. (See Exercise VIII.7.4.) If $n > 1$ we have $\int_\Gamma (z - z_0)^{-n} dz = 0$ because the integrand has a primitive, and if $n = 1$ the same conclusion holds because all points having a nonzero index with respect to Γ lie in K.

By Runge's theorem, there is a sequence $(g_k)_{k=1}^\infty$ of rational functions with poles in $\mathbf{C}\backslash K$ that converges uniformly to f on K, and so in particular on Γ. Then

$$\int_\Gamma f(z)dz = \lim_{k\to\infty} \int_\Gamma g_k(z)dz.$$

Since, as shown above, $\int_\Gamma g_k(z)dz = 0$ for each k, it follows that $\int_\Gamma f(z)dz = 0$, as desired.

IX.17. Sharpened Form of Runge's Theorem

Let K be a compact subset of $\mathbf{C}$, and let S be a subset of $\overline{\mathbf{C}}\backslash K$ that contains at least one point in each connected component of $\overline{\mathbf{C}}\backslash K$. Then any function holomorphic in an open set containing K can be uniformly approximated on K by rational functions whose poles lie in S.

In particular, if $\mathbf{C}\backslash K$ is connected we can take S to be the singleton $\{\infty\}$, giving in this case the conclusion that any function holomorphic in an open set containing K can be uniformly approximated on K by polynomials.

To prove the sharpened form of Runge's theorem, it will suffice to show that any rational function whose poles lie in $\mathbf{C}\backslash K$ can be uniformly approximated on K by rational functions whose poles lie in S. Because any rational function can be written as a sum of rational functions each having only one pole (by Exercise VIII.7.4), the desired conclusion is a consequence of the following assertion: If the points z_0 of $\mathbf{C}$ and z_0' of $\overline{\mathbf{C}}$ lie in the same connected component of $\overline{\mathbf{C}}\backslash K$, then any rational function whose only pole is z_0 can be uniformly approximated on K by rational functions whose only pole is z_0'.

We treat first the case $z_0' \neq \infty$. In that case z_0 and z_0' lie in the same connected component of $\mathbf{C}\backslash K$, so there is a polygonal path γ in that component whose initial point is z_0 and whose terminal point is z_0'. (We are using here a basic property of connected open subsets of the plane. The reader unfamiliar with it can establish it as Exercise IX.17.1 below.) On γ we can find a finite sequence of points $z_0, z_1, \ldots, z_q = z_0'$ such that, for each j, the distance $|z_{j-1} - z_j|$ is less than the distance from γ to K. Clearly, it will suffice to show that, for each j, any rational function whose only pole is z_{j-1} can be uniformly approximated on K by rational functions whose only pole is z_j. But a rational function whose only pole is z_{j-1} has a Laurent series representation with center z_j that converges locally uniformly to the function in the region $|z - z_j| > |z_{j-1} - z_j|$, a region containing K (since $|z_{j-1} - z_j| < \operatorname{dist}(\gamma, K) \leq \operatorname{dist}(z_j, K)$). Because the given rational function has a removable singularity at ∞, the Laurent series contains no positive powers of $z - z_j$. Its partial sums are therefore rational functions having z_j as their only pole, and they converge uniformly to the given rational function on K, as desired.

It remains to treat the case $z_0' = \infty$. In that case z_0 lies in the unbounded component of $\mathbf{C}\backslash K$. By what has just been proved, we can then assume with no loss of generality that K is contained in the disk $|z| < |z_0|$. In that case, if a rational function has z_0 as its only pole, its Taylor series centered at 0 converges uniformly to it on K, and thus the function can be uniformly approximated on K by polynomials. This completes the proof.

Exercise IX.17.1. Prove that if two points lie in the same connected open subset G of $\mathbf{C}$, then the points can be joined by a polygonal path that lies in G.

Exercise IX.17.2. Let G be an open subset of $\mathbf{C}$ and S a subset of $\overline{\mathbf{C}}\backslash G$ that contains at least one point in each connected component of $\overline{\mathbf{C}}\backslash G$. Let f be a holomorphic function in G. Prove that there is a sequence of rational functions whose poles lie in S that converges locally uniformly to f in G.

Exercise IX.17.3. Prove that there is a sequence $(p_k)_{k=1}^{\infty}$ of polynomials such that

$$\lim_{k\to\infty} p_k(z) \;=\; \begin{cases} 1, & \operatorname{Re} z > 0 \\ 0, & \operatorname{Re} z = 0 \\ -1, & \operatorname{Re} z < 0. \end{cases}$$

Chapter X

Further Development of Basic Complex Function Theory

The chapter begins with a topological notion, that of a simply connected domain, followed by several characterizations of simply connected domains. The term "domain" for us will mean a nonempty, connected, open subset of $\mathbf{C}$, a usage common in complex analysis. The notion of simple connectivity is a general topological one. The definition used here for plane domains, while arguably the most convenient for complex analysis, differs from the general topological definition, which involves homotopy. The equivalence of the two definitions will be established at the end of the chapter.

Next, various consequences of Cauchy's theorem will be obtained, notably, the residue theorem together with examples of its use in evaluating definite integrals, and the argument principle, which gives an integral expression for the number of zeros of a holomorphic function in a region. The chapter culminates with a proof of the famous Riemann mapping theorem.

X.1. Simply Connected Domains

As noted above, the term "domain" is commonly used in complex function theory to designate a nonempty, connected, open set. We shall henceforth use the term in this sense. A domain G in $\mathbf{C}$ is said to be simply connected if $\overline{\mathbf{C}}\backslash G$, its extended complement, is connected.

The reader will easily verify that every convex domain—in particular, every open disk and every open half-plane—is simply connected. The simplest example of a domain that is not simply connected is $\mathbf{C}\backslash\{0\}$. A bounded domain G is simply connected if and only if $\mathbf{C}\backslash G$, its ordinary complement, is connected. However, the ordinary complement of an unbounded simply connected domain can fail to be connected. The strip $0 < \operatorname{Im} z < 1$ furnishes a simple example.

X.2. Winding Number Criterion

For the domain G in $\mathbf{C}$ to be simply connected, it is necessary and sufficient that every contour in G have winding number 0 about every point in $\overline{\mathbf{C}}\backslash G$.

To establish the necessity, assume G is simply connected, and let Γ be a contour in G. Since $\overline{\mathbf{C}}\backslash G$ is connected and contained in $\overline{\mathbf{C}}\backslash\Gamma$, it is contained in a single connected component of $\overline{\mathbf{C}}\backslash\Gamma$, necessarily the unbounded component. Hence $\mathbf{C}\backslash G$ is contained in the unbounded component of $\mathbf{C}\backslash\Gamma$. Since Γ has winding number 0 about every point in the unbounded component of $\mathbf{C}\backslash\Gamma$, the desired conclusion follows.

For the other direction, assume that G is not simply connected. Then there is a decomposition $\overline{\mathbf{C}}\backslash G = K \cup L$, where K and L are closed, disjoint, and nonempty, say with ∞ in L. Let $H = \overline{\mathbf{C}}\backslash L$. Then H is an open subset of $\mathbf{C}$, and H contains K. By the separation lemma IX.8, there is a simple contour in $H\backslash K$ $(= G)$ that has winding number 1 about each point of K. This establishes the sufficiency of the condition.

Exercise X.2.1. Let G be an open subset of $\mathbf{C}$. Prove that $\overline{\mathbf{C}}\backslash G$ is connected if and only if every connected component of G is simply connected.

Exercise X.2.2. Let $(G_n)_{n=1}^{\infty}$ be a decreasing sequence of simply connected domains such that the interior of $\bigcap_1^\infty G_n$ is nonempty. ("Decreasing" means $G_{n+1} \subset G_n$ for all n.) Prove that every connected component of the interior of $\bigcap_1^\infty G_n$ is simply connected.

Exercise X.2.3. Prove that the union of an increasing sequence of simply connected domains is simply connected.

X.3. Cauchy's Theorem for Simply Connected Domains

Cauchy's theorem states that a holomorphic function in an open set G has integral 0 around any contour in G that has winding number 0 about every point of $\mathbf{C}\backslash G$. By X.2, if G is a simply connected domain, the preceding winding number condition holds automatically. Hence: *If f is a holomorphic function in the simply connected domain G, then $\int_\Gamma f(z)dz = 0$ for every contour Γ in G.*

X.4. Existence of Primitives

For the domain G in $\mathbf{C}$ to be simply connected, it is necessary and sufficient that every holomorphic function in G have a primitive.

That holomorphic functions in convex domains have primitives was proved, in Section VII.2, as the main step in the proof of Cauchy's theorem for such domains. The argument depended on Cauchy's theorem for a triangle. Now that we have the general Cauchy theorem, we can use the same argument in a general simply connected domain: Suppose G is simply connected, and let f be a holomorphic function in G. Fix a point z_0 in G, and define the function g in G by $g(z) = \int_\gamma f(\zeta)d\zeta$, where γ is any piecewise-C^1 curve in G that starts at z_0 and ends at z. Cauchy's theorem implies that the definition of g makes sense; in other words, the integral above has the same value for any two curves γ satisfying the stated conditions. One can now prove exactly as in Section VII.2 that g is holomorphic and $g' = f$. This establishes the necessity of the condition.

For the other direction, assume that G is not simply connected. Then, by X.2, there is a contour Γ in G such that $\displaystyle\int_\Gamma \frac{1}{z - z_0}dz \neq 0$ for some point z_0 in $\mathbf{C}\backslash G$. The function $\dfrac{1}{z - z_0}$ then has no primitive in G. This establishes the sufficiency of the condition.

Exercise X.4.1. Let G be a simply connected domain, f a holomorphic function in G, and c a positive number. Prove that each connected component of the set $\{z \in G : |f(z)| < c\}$ is simply connected.

X.5. Existence of Logarithms

Let G be a simply connected domain and f a nowhere vanishing holomorphic function in G. Then there is a branch of $\log f$ *in G.*

To prove this, fix a point z_0 in G, and let a be any value of $\log f(z_0)$. By X.4, the function $\dfrac{f'}{f}$ has a primitive, g, in G, and we may obviously assume that $g(z_0) = a$. We have

$$(fe^{-g})' = f'e^{-g} - g'fe^{-g} = 0,$$

implying that e^g is a constant times f. Since f and e^g have the same value at z_0, the constant is 1. Thus g is a branch of $\log f$ in G, as desired.

Exercise X.5.1. Prove the converse of X.5: If the domain G in $\mathbf{C}$ has the property that, for every nowhere vanishing holomorphic function f in G, there is a branch of $\log f$ in G, then G is simply connected.

Exercise X.5.2. Let G be a domain in $\mathbf{C}$ and f a nowhere vanishing holomorphic function in G. Prove that there is a branch of $\log f$ in G if and only if there is a branch of $f^{1/n}$ in G for every positive integer n.

X.6. Existence of Harmonic Conjugates

Let G be a simply connected domain in $\mathbf{C}$ and u a real-valued harmonic function in G. Then u has a harmonic conjugate in G, unique to within addition of a real constant.

This is proved for convex domains in Section VII.18. The argument there can now be applied word-for-word to the more general case of a simply connected domain.

Exercise X.6.1. Let the domain G in $\mathbf{C}$ be multiply connected (i.e., not simply connected). Prove that there is a real-valued harmonic function u in G that has no harmonic conjugate in G.

X.7. Simple Connectivity and Homotopy

If G is a domain in $\mathbf{C}$, and if every closed curve in G is null homotopic in G, then G is simply connected.

A closed curve in G is said to be null homotopic in G if it is homotopic in G to a constant curve, i.e., a curve whose range is a singleton. In that case one says the curve can be shrunk in G to a point.

To prove the statement we note that, if G has the given property, then by IX.13 every closed curve in G has winding number 0 about every point in $\mathbf{C}\backslash G$, which by X.2 implies that G is simply connected.

Our definition of simple connectivity for a domain G in the plane is deficient from a topological perspective because it is not intrinsic to G: it refers to the complement $\mathbf{C}\backslash G$. Not apparent from our definition is whether simple connectivity is a topological property of a domain. Put in other terms, if the domains G_1 and G_2 are homeomorphic (that is, in one-to-one correspondence under a bicontinuous map), and if G_1 is simply connected, must G_2 be simply connected?

According to the general definition, a topological space is called simply connected if it is pathwise connected (any two points in the space can be joined by a curve in the space) and every closed curve in the space is null homotopic. Simple connectivity in this sense is clearly a topological property. According to the proposition at the beginning of this section, for a domain in $\mathbf{C}$, simple connectivity in the homotopic sense implies simple connectivity in our sense. The converse implication is given in Section X.21. Thus, the definition of simple connectivity we use is equivalent to the standard definition, in the context of plane domains.

The reader wishing to learn more about these matters is referred to the books of R. B. Burckel and M. H. A. Newman in the reference list.

X.8. The Residue Theorem

Let G be a domain in $\mathbf{C}$ and Γ a contour contained with its interior in G. Let the function f be holomorphic in G except for isolated singularities at the points $z_1, \ldots, z_p$, none of which lies on Γ. Then

$$\int_\Gamma f(z)dz = 2\pi i \sum_{k=1}^{p} \operatorname{ind}_\Gamma(z_k) \cdot \operatorname{res}_{z_k}(f).$$

To prove this we choose, for each singularity z_k, a counterclockwise oriented circle C_k with center z_k, contained along with its interior in G, and disjoint along with its interior from Γ and from all of the other circles. Let $n_k = \operatorname{ind}_\Gamma(z_k),\ k = 1, \ldots, p$. We consider the domain

$$G_1 = G \backslash \{z_1, \ldots z_p\}$$

and the contour

$$\Gamma_1 = \Gamma - \sum_{k=1}^{p} n_k C_k.$$

The function f is holomorphic in G_1, and the contour Γ_1 lies in G_1 and has winding number 0 about each point in $\mathbf{C} \backslash G_1$. By Cauchy's theorem, $\int_{\Gamma_1} f(z)dz = 0$, which can be written

$$\int_\Gamma f(z)dz = \sum_{k=1}^{p} n_k \int_{C_k} f(z)dz.$$

Since $\int_{C_k} f(z)dz = 2\pi i \operatorname{res}_{z_k}(f)$ for each k, the desired equality follows.

Exercise X.8.1. Let C be the unit circle with the counterclockwise orientation, and let a and b be points in the interior of C. Evaluate the integral

$$\int_C \frac{z^k}{(z-a)(z-b)^2} dz,$$

where k is a positive integer.

Exercise X.8.2. Let p and q be polynomials such that $\deg q > 1 + \deg p$. Prove that the sum of the residues of the rational function $\frac{p}{q}$, taken over all of its poles in $\mathbf{C}$, is 0.

Exercise X.8.3. Evaluate the integral

$$\int_C (z-2)^{-1}(2z+1)^{-2}(3z-1)^{-3}dz,$$

where C is the unit circle with the counterclockwise orientation. (Suggestion: Try to use the preceding exercise.)

Exercise X.8.4. Evaluate the integrals

$$\int_C \frac{\sin z}{2z^2-5z+2}dz, \qquad \int_C \frac{2z^2-5z+2}{\sin z}dz$$

where C is the unit circle with the counterclockwise orientation.

Exercise X.8.5. Extend the residue theorem to the case where the function f has countably many isolated singularities in the domain G.

X.9. Cauchy's Formula

Let G be a domain in $\mathbf{C}$, f a holomorphic function in G, and Γ a simple contour contained with its interior in G. Then

$$f(z_0) = \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-z_0}dz, \qquad z_0 \in \operatorname{int}\Gamma.$$

This is an immediate consequence of the residue theorem: under the given hypotheses, the integrand in the preceding integral is holomorphic in $G\backslash\{z_0\}$, and its residue at z_0 is $f(z_0)$.

Cauchy's formula for the n-th derivative ($n = 1, 2, 3, \dots$) is a consequence: under the given hypothesis:

$$f^{(n)}(z_0) = \frac{n!}{2\pi i}\int_\Gamma \frac{f(z)}{(z-z_0)^{n+1}}dz, \qquad z_0 \in \operatorname{int}\Gamma.$$

The deduction is contained in the discussion of Cauchy integrals in Section VII.7.

X.10. More Definite Integrals

The exploitation of complex integration to evaluate improper Riemann integrals has been illustrated in Sections VI.12 and VII.4. Now that we have further developed the theory, we can handle many additional integrals. The residue theorem is an especially powerful tool for this purpose.

Example 1. $\displaystyle\int_0^\infty \frac{\sin x}{x}dx.$

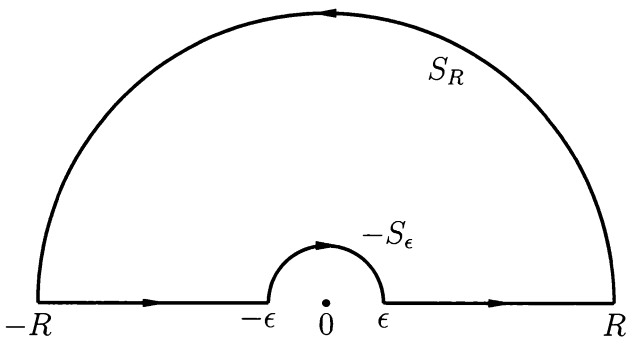

Figure 8. The contour $\Gamma_{\epsilon,R}$ for Example 1.

This is an important integral in Fourier analysis. It is not obvious that the integral converges. Although the integrand is well behaved at the left endpoint of the interval of integration, there is potential trouble at the far end, because $\int_0^\infty \left|\frac{\sin x}{x}\right| dx$ diverges (Exercise X.10.1 below). Nevertheless, as we shall prove on the basis of Cauchy's theorem, the given integral does converge.

For $r > 0$, let S_r denote the semicircle in the upper half-plane with center 0 and radius r, oriented counterclockwise. For $0 < \epsilon < R$, let $\Gamma_{\epsilon,R}$ denote the closed curve consisting of the interval $[\epsilon, R]$ followed by the semicircle S_R followed by the interval $[-R, -\epsilon]$ followed by the semicircle $-S_\epsilon$ (see Figure 8). The function $\frac{e^{iz}}{z}$, whose imaginary part on the real axis equals $\frac{\sin x}{x}$, is holomorphic in $\mathbf{C}\backslash\{0\}$, a domain containing $\Gamma_{\epsilon,R}$ and its interior. By Cauchy's theorem,

$$\int_{\Gamma_{\epsilon,R}} \frac{e^{iz}}{z} dz = 0.$$

(We are using here the general Cauchy theorem. By an argument like that in Section VII.5, the perverse reader could draw the same conclusion using only Cauchy's theorem for convex domains. Now that we have the general Cauchy theorem, such contortions are unnecessary.)

On the real axis, the real part of $\frac{e^{ix}}{x}$ is $\frac{\cos x}{x}$, an odd function (undefined at 0), and the imaginary part, as already noted, is $\frac{\sin x}{x}$, an even function. The preceding equality therefore implies that

$$(*) \qquad 2i\int_\epsilon^R \frac{\sin x}{x} dx = -\int_{S_R} \frac{e^{iz}}{z} dz + \int_{S_\epsilon} \frac{e^{iz}}{z} dz.$$

We shall now take the limit as $R \to \infty$ and $\epsilon \to 0$.

Consider first the integral over S_R. We shall prove that it tends to 0 as $R \to \infty$. This is intuitively plausible, because, in the upper half-plane, $|e^{iz}|$ is very small except in the vicinity of the real axis. Introducing the parametrization $t \mapsto Re^{it}$ $(0 \leq t \leq \pi)$ of S_R, we can write

$$\int_{S_R} \frac{e^{iz}}{z} dz = i \int_0^{\pi} e^{iRe^{it}} dt,$$

and so

$$\begin{aligned} \left| \int_{S_R} \frac{e^{iz}}{z} dz \right| &\leq \int_0^{\pi} \left| e^{iRe^{it}} \right| dt \\ &= \int_0^{\pi} e^{-R \sin t} dt \\ &= 2 \int_0^{\pi/2} e^{-R \sin t} dt. \end{aligned}$$

Using the inequality $\sin t \geq 2t/\pi$ $(0 \leq t \leq \pi/2)$, we obtain

$$\begin{aligned} \left| \int_{S_R} \frac{e^{iz}}{z} dz \right| &\leq 2 \int_0^{\pi/2} e^{-2Rt/\pi} dt \\ &= \frac{\pi}{R}\left(1 - e^{-R}\right). \end{aligned}$$

Because the right side tends to 0 as $R \to \infty$, we can conclude that

$$\lim_{R \to \infty} \int_{S_R} \frac{e^{iz}}{z} dz = 0,$$

as desired.

Consider now the integral over S_ϵ. The function $f(z) = \dfrac{e^{iz} - 1}{z}$ has a removable singularity at the origin; in particular, it stays bounded near the origin. From this, one easily concludes that

$$\lim_{\epsilon \to 0} \int_{S_\epsilon} f(z) dz = 0.$$

Since $\dfrac{e^{iz}}{z} = \dfrac{1}{z} + f(z)$ and

$$\int_{S_\epsilon} \frac{1}{z} dz = \pi i$$

for all ϵ (by a straightforward calculation), we obtain

$$\lim_{\epsilon \to 0} \int_{S_\epsilon} \frac{e^{iz}}{z} dz = \pi i.$$

Combining the two limits just found with the equality $(*)$, we obtain

$$\lim_{\substack{\epsilon\to 0\\ R\to\infty}} \int_{\epsilon}^{R} \frac{\sin x}{x}dx = \frac{\pi}{2}.$$

In other words,

$$\int_0^{\infty} \frac{\sin x}{x}dx = \frac{\pi}{2}.$$

Example 2. $\displaystyle\int_{-\infty}^{\infty} \frac{\cos x}{1+x^4}dx.$

In this example we can see at the outset, using a comparison test, that the integral converges. (Details are left to the reader; prior knowledge of convergence will not be used below.) As in the preceding example, we make use of the function e^{iz}, whose absolute value in the upper half-plane is bounded by 1. The integrand in our integral coincides on the real axis with the real part of the function $\dfrac{e^{iz}}{1+z^4}$. That function is holomorphic in $\mathbf{C}$ except for simple poles at the points $e^{\pi in/4}$, $n = 1, 3, 5, 7$ (the fourth roots of -1).

For $R > 1$, let Γ_R denote the closed curve consisting of the interval $[-R, R]$ followed by the semicircle S_R (defined as in the preceding example). The singularities of the function $\dfrac{e^{iz}}{1+z^4}$ in the interior of Γ_R are the points $z_1 = e^{\pi i/4}$ and $z_2 = e^{3\pi i/4}$. By the residue theorem,

$$\int_{\Gamma_R} \frac{e^{iz}}{1+z^4}dz = 2\pi i\left[\operatorname{res}_{z=z_1}\left(\frac{e^{iz}}{1+z^4}\right) + \operatorname{res}_{z=z_2}\left(\frac{e^{iz}}{1+z^4}\right)\right].$$

To evaluate the residues we use the result of Exercise VIII.12.1: if the functions g and h are holomorphic in an open set containing the point z_0 and h has a simple zero at z_0, then $\operatorname{res}_{z_0} \dfrac{g}{h} = \dfrac{g(z_0)}{h'(z_0)}$. We thus have

$$\begin{aligned}
\operatorname{res}_{z=z_1}\left(\frac{e^{iz}}{1+z^4}\right) &= \frac{\exp(iz_1)}{4z_1^3} \\
&= \frac{\exp\left(-\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right)}{4\left(-\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right)} \\
&= \frac{1}{4}\left(-\frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}\right)e^{-1/\sqrt{2}}\left(\cos\frac{1}{\sqrt{2}} + i\sin\frac{1}{\sqrt{2}}\right),
\end{aligned}$$

$$\begin{aligned}
\operatorname{res}_{z=z_2}\left(\frac{e^{iz}}{1+z^4}\right) &= \frac{\exp(iz_2)}{4z_2^3} \\
&= \frac{\exp\left(-\frac{1}{\sqrt{2}}-\frac{i}{\sqrt{2}}\right)}{4\left(\frac{1}{\sqrt{2}}+\frac{i}{\sqrt{2}}\right)} \\
&= \frac{1}{4}\left(\frac{1}{\sqrt{2}}-\frac{i}{\sqrt{2}}\right)e^{-1/\sqrt{2}}\left(\cos\frac{1}{\sqrt{2}}-i\sin\frac{1}{\sqrt{2}}\right).
\end{aligned}$$

The residue at z_2 is the negative of the complex conjugate of the residue at z_1, so that the sum of the two residues equals $2i$ times the imaginary part of the residue at z_1. The sum of the two residues therefore equals

$$\frac{i\sqrt{2}}{4}e^{-1/\sqrt{2}}\left(-\cos\frac{1}{\sqrt{2}}-\sin\frac{1}{\sqrt{2}}\right).$$

It follows that

$$\int_{\Gamma_R}\frac{e^{iz}}{1+z^4}dz = \frac{\pi\sqrt{2}}{2}e^{-1/\sqrt{2}}\left(\cos\frac{1}{\sqrt{2}}+\sin\frac{1}{\sqrt{2}}\right).$$

On the real axis, the imaginary part of $\dfrac{e^{ix}}{1+x^4}$ is an odd function. Thus,

$$\int_{\Gamma_R}\frac{e^{iz}}{1+z^4}dz = \int_{-R}^{R}\frac{\cos x}{1+x^4}dx + \int_{S_R}\frac{e^{iz}}{1+z^4}dz.$$

This in combination with the preceding equality gives

$$\int_{-R}^{R}\frac{\cos x}{1+x^4}dx = \frac{\pi\sqrt{2}}{2}e^{-1/\sqrt{2}}\left(\cos\frac{1}{\sqrt{2}}+\sin\frac{1}{\sqrt{2}}\right) - \int_{S_R}\frac{e^{iz}}{1+z^4}dz.$$

We now take the limit as $R\to\infty$. The absolute value of the integrand in the integral over S_R is bounded by $\dfrac{1}{R^4-1}$, so the integral itself is bounded in absolute value by $\dfrac{\pi R}{R^4-1}$, which tends to 0 as $R\to\infty$. We can conclude that

$$\int_{-\infty}^{\infty}\frac{\cos x}{1+x^4}dx = \frac{\pi\sqrt{2}}{2}e^{-1/\sqrt{2}}\left(\cos\frac{1}{\sqrt{2}}+\sin\frac{1}{\sqrt{2}}\right).$$

Example 3. $\displaystyle\int_0^{\infty}\frac{1}{1+x^a}dx, \quad a>1.$

As in the last example, one can see at the outset, by means of a comparison test, that the integral converges, although that knowledge will not be required below. In the domain $G=\{re^{i\theta} : r>0, \left|\theta-\frac{\pi}{a}\right|<\pi\}$ we consider the branch of the function z^a that takes the value 1 at the point $z=1$. We

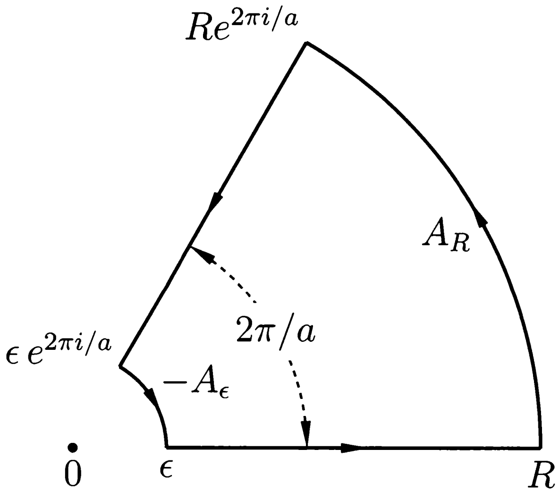

Figure 9. The contour $\Gamma_{\epsilon,R}$ for Example 3.

denote this function simply by z^a. The function $\dfrac{1}{1+z^a}$ is holomorphic in G except for a simple pole at the point $z_0 = e^{\pi i/a}$.

For $r > 0$, let A_r denote the circular arc $\left\{re^{i\theta} : 0 < \theta < \frac{2\pi}{a}\right\}$, oriented counterclockwise. For $0 < \epsilon < 1 < R$, let $\Gamma_{\epsilon,R}$ denote the closed curve consisting of the interval $[\epsilon, R]$ followed by the arc A_R followed by the segment $\left[Re^{2\pi i/a}, \epsilon e^{2\pi i/a}\right]$ followed by the arc $-A_\epsilon$ (see Figure 9). By the residue theorem,

$$\int_{\Gamma_{\epsilon,R}} \frac{1}{1+z^a}dz \;=\; 2\pi i \;\mathrm{res}_{z=z_0}\left(\frac{1}{1+z^a}\right).$$

We can compute the residue by the same method we used in the preceding example. We obtain

$$\mathrm{res}_{z=z_0}\left(\frac{1}{1+z^a}\right) \;=\; \frac{1}{az_0^{a-1}} \;=\; \frac{-e^{\pi i/a}}{a}.$$

Thus

$$\int_{\Gamma_{\epsilon,R}} \frac{1}{1+z^a}dz \;=\; \frac{-2\pi i e^{\pi i/a}}{a}.$$

Now, the integral of $\dfrac{1}{1+z^a}$ over the interval $[\epsilon, R]$ equals $\displaystyle\int_\epsilon^R \frac{1}{1+x^a}dx$, and the integral of the same function along the segment $\left[\epsilon e^{2\pi i/a}, Re^{2\pi i/a}\right]$ equals $e^{2\pi i/a}\displaystyle\int_\epsilon^R \frac{1}{1+x^a}dx$ (as one sees by using the parametrization $t \mapsto$

$te^{2\pi i/a}$ $(\epsilon \le t \le R)$). We can therefore rewrite the preceding equality as

$$\left(1 - e^{2\pi i/a}\right) \int_\epsilon^R \frac{1}{1+x^a} dx = \frac{-2\pi i e^{\pi i/a}}{a} + \int_{A_\epsilon} \frac{1}{1+z^a} dz - \int_{A_R} \frac{1}{1+z^a} dz.$$

We now take the limit as $\epsilon \to 0$ and $R \to \infty$. The integral over A_ϵ clearly tends to 0 (since the integrand stays bounded). In the integral over A_R, the integrand is bounded in absolute value by $\frac{1}{R^a - 1}$, so the integral itself is bounded in absolute value by $\frac{2\pi R}{a(R^a - 1)}$, which tends to 0 as $R \to \infty$. In the limit we thus obtain

$$\left(1 - e^{2\pi i/a}\right) \int_0^\infty \frac{1}{1+x^a} dx = \frac{-2\pi i e^{\pi i/a}}{a},$$

giving (after a short calculation)

$$\int_0^\infty \frac{1}{1+x^a} dx = \frac{\pi}{a \sin \frac{\pi}{a}}.$$

Exercise X.10.1. Prove that the integral $\int_0^\infty \left|\frac{\sin x}{x}\right| dx$ diverges.

Exercise X.10.2. Derive the formula

$$\int_{-\infty}^\infty \frac{\cos x}{\cosh x} dx = \frac{\pi}{\cosh \frac{\pi}{2}}$$

by integrating the function $\frac{e^{iz}}{\cosh z}$ around the rectangle with vertices $-R$, R, $R + \pi i$, $-R + \pi i$, and letting $R \to \infty$.

Exercise X.10.3. Evaluate $\int_0^\infty \frac{\sin^2 x}{x^2} dx$ by integrating the function $\frac{1 - e^{2iz}}{z^2}$ around the curves $\Gamma_{\epsilon,R}$ used in Example 1, and taking the limit as $\epsilon \to 0$ and $R \to \infty$.

Exercise X.10.4. Derive the formula

$$\int_0^\infty \frac{\cos ax}{(x^2 + b^2)^2} dx = \frac{\pi}{4b^3}(1 + ab)e^{-ab}, \qquad a \ge 0,\ b > 0.$$

Exercise X.10.5. Derive the formula

$$\int_0^\infty \frac{x^{a-1}}{1+x^2} dx = \frac{\pi}{2 \sin\left(\frac{\pi a}{2}\right)}, \qquad 0 < a < 2.$$

Exercise X.10.6. Evaluate

$$\int_0^{2\pi} \frac{\cos\theta}{a - \cos\theta} d\theta, \qquad a > 1.$$

Exercise X.10.7. Let the function $f = u + iv$ be holomorphic in a domain containing the closed unit disk. Derive the relations

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}\frac{e^{i\theta}+z}{e^{i\theta}-z}f(e^{i\theta})d\theta &= 2f(z)-f(0), \qquad |z|<1,\\ \frac{1}{2\pi}\int_0^{2\pi}\frac{e^{i\theta}+z}{e^{i\theta}-z}\overline{f(e^{i\theta})}d\theta &= \overline{f(0)}, \qquad |z|<1.\end{aligned}$$

From these deduce Herglotz's formula,

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi}\frac{e^{i\theta}+z}{e^{i\theta}-z}u(e^{i\theta})d\theta + iv(0), \qquad |z|<1,$$

and Poisson's formula,

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi}\frac{1-|z|^2}{|e^{i\theta}-z|^2}u(e^{i\theta})d\theta, \qquad |z|<1.$$

(These formulas are of central importance in more advanced function theory.)

X.11. The Argument Principle

Let G be a domain in $\mathbf{C}$ and Γ a simple contour contained with its interior in G. Let f be a holomorphic function in G without zeros on Γ. Then the number of zeros of f in int Γ, taking into account multiplicities, equals

$$\frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}dz.$$

From the discussion in Section IX.3 we know that

$$\frac{1}{i}\int_\Gamma \frac{f'(z)}{f(z)}dz$$

equals the increment experienced by arg $f(z)$ as z makes one circuit of Γ. The argument principle thus states that each zero of f in int Γ accounts for 2π of this increment. Multiple zeros, as indicated in the statement of the principle, are counted as many times as required by their multiplicities.

To establish the argument principle we first note that f can have only finitely many zeros in int Γ. In fact, the set ext Γ is open and contains both $\mathbf{C}\backslash G$ and the unbounded connected component of $\mathbf{C}\backslash\Gamma$. The set $\mathbf{C}\backslash$ ext Γ is thus a closed and bounded (i.e., compact) subset of G. Since the zero set of f has no limit points in G (see Section VII.13), there can be only finitely many zeros of f in $\mathbf{C}\backslash$ ext Γ, and thus there are only finitely many zeros of f in int Γ, as asserted.

Suppose $z_1, \ldots, z_p$ are the zeros of f in int Γ, and let $m_1, \ldots, m_p$ be their respective orders. Let G_1 be the domain obtained by removing from G the zeros of f in ext Γ. Then Γ and its interior are contained in G_1, where the function f'/f is holomorphic except for simple poles at the points $z_1, \ldots, z_p$. By the residue theorem,

$$\int_\Gamma \frac{f'(z)}{f(z)}dz = 2\pi i \sum_{k=1}^{p} res_{z_k}\left(\frac{f'}{f}\right).$$

A simple calculation (requested as Exercise VIII.12.3) shows that

$$res_{z_k}\left(\frac{f'}{f}\right) = m_k,$$

and the desired equality follows.

Exercise X.11.1. Evaluate

$$\frac{1}{2\pi i}\int_C \frac{z^{n-1}}{3z^n - 1}dz,$$

where n is a positive integer, and C is the unit circle with the counterclockwise orientation.

Exercise X.11.2. (Argument principle for meromorphic functions.) Let G be a domain in $\mathbf{C}$ and Γ a simple contour contained with its interior in G. Let the function f be holomorphic in G except for isolated poles. (Such an f is said to be meromorphic in G.) Prove that if f has neither zeros nor poles on Γ, then

$$\frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}dz$$

equals the number of zeros that f has in int Γ minus the number of poles that f has in int Γ, taking account of multiplicities.

Exercise X.11.3. Let the function f be holomorphic in a domain containing the closed unit disk. Prove that the increment in arg f around the unit circle equals 2π times the total number of zeros of f in the closed unit disk, provided zeros on the unit circle are counted with one-half their multiplicities. (Part of the problem is to make a reasonable definition of the increment in arg f.)

X.12. Rouché's Theorem

Let G be a domain in $\mathbf{C}$, K a compact subset of G, and f and g holomorphic functions in G such that $|f(z) - g(z)| < |f(z)|$ for every point z in the boundary of K. Then f and g have the same number of zeros in the interior of K, taking into account multiplicities.

Notice that the hypotheses imply that neither f nor g vanishes on the boundary of K. Roughly, the theorem states that the number of zeros a holomorphic function has in a compact set is stable under small perturbations of the function, provided zeros on the boundary of the set are absent.

To prove Rouché's theorem, let G_1 denote the union of K with the set of points z in G such that $|f(z) - g(z)| < |f(z)|$. Then G_1 is an open set containing K. By the separation lemma IX.8, there is a simple contour Γ in $G_1 \backslash K$ such that $K \subset \operatorname{int} \Gamma \subset G_1$.

For $0 \leq t \leq 1$ let $f_t = (1-t)f + tg$, so that $f_0 = f$ and $f_1 = g$. Let $m(t)$ denote the number of zeros of f_t in the interior of K. From the expression $f_t = f - t(f-g)$ it is clear that f_t is nonzero wherever the inequality $|f - g| < |f|$ holds. Hence, the only zeros of f_t in G_1 are the zeros in the interior of K. Therefore, $m(t)$ equals the number of zeros of f_t in $\operatorname{int} \Gamma$, and by the argument principle we have

$$m(t) = \frac{1}{2\pi i} \int_\Gamma \frac{f_t'(z)}{f_t(z)} dz.$$

Finally, standard estimates show that the integral on the right side of the preceding equality is a continuous function of t (Exercise X.12.1 below). Since $m(t)$ depends continuously on t and is integer valued, it is constant. In particular, $m(1) = m(0)$, the desired conclusion.

Exercise X.12.1. Supply the argument, omitted in the preceding proof, to establish the continuity of the function $m(t)$.

Exercise X.12.2. How many zeros, taking into account multiplicities, does the polynomial $z^7 + 4z^4 + z^3 + 1$ have in the regions $|z| < 1$ and $1 < |z| < 2$?

Exercise X.12.3. Prove that, for any positive number ϵ, the function $\dfrac{1}{z+i} + \sin z$ has infinitely many zeros in the strip $|\operatorname{Im} z| < \epsilon$.

Exercise X.12.4. Let f be a holomorphic map of the open unit disk into itself whose range lies in a compact subset of the open disk. Prove that f has a unique fixed point.

Exercise* X.12.5. (Hurwitz's theorem.) Let the sequence $(f_n)_{n=1}^\infty$ of holomorphic functions in the domain G converge locally uniformly in G to the nonconstant function f. Prove that if f has at least m zeros in G, then all but finitely many of the functions f_n have at least m zeros in G. Conclude, as a corollary, that if every f_n is univalent then f is univalent.

Exercise X.12.6. Prove that all of the zeros of the polynomial $z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0$ lie in the disk with center 0 and radius

$$\sqrt{1 + |c_{n-1}|^2 + \cdots + |c_1|^2 + |c_0|^2}.$$

Exercise X.12.7. Prove that, if $1 < a < \infty$, the function $z + a - e^z$ has only one zero in the half-plane $\operatorname{Re} z < 0$, and this zero is on the real axis.

Exercise X.12.8. Prove that, if $0 < |a| < 1$ and n is a positive integer, then the equation $(z-1)^n e^z = a$ has exactly n roots, each of multiplicity 1, in the half-plane $\operatorname{Re} z > 0$. Prove that if $|a| \leq 2^{-n}$ the roots are all in the disk $|z-1| < 1/2$.

X.13. The Local Mapping Theorem

Let f be a nonconstant holomorphic function in the domain G. Let z_0 be a point of G and let $w_0 = f(z_0)$. Let m be the multiplicity with which f assumes the value w_0 at z_0 (i.e., m is the order of the zero of $f - w_0$ at z_0). For every sufficiently small positive number δ, there is a positive number ϵ such that every value w satisfying $0 < |w - w_0| < \epsilon$ is assumed by f at exactly m distinct points in the punctured disk $0 < |z - z_0| < \delta$, with multiplicity 1 at each of those points.

The theorem says that, near the point z_0, the function f behaves qualitatively like the polynomial $(z-z_0)^m + w_0$. To prove the theorem we choose a positive number δ_0 such that the closed disk $|z - z_0| \leq \delta_0$ is contained in G, and such that neither $f - w_0$ nor f' vanishes for $0 < |z - z_0| \leq \delta_0$; such a choice is possible because the zeros of a nonconstant holomorphic function are isolated. We shall show that the desired conclusion holds for every positive number δ not exceeding δ_0.

Let δ be as described, and let K denote the closed disk $|z - z_0| \leq \delta$. Let ϵ denote the minimum of $|f(z) - w_0|$ for z on the boundary of K; it is positive because f does not assume the value w_0 in K except at z_0. Fix w such that $0 < |w - w_0| < \epsilon$. Then for z in the boundary of K we have

$$\big|(f(z) - w_0) - (f(z) - w)\big| \;=\; |w - w_0| \;<\; \epsilon \;\leq\; |f(z) - w_0|.$$

It therefore follows by Rouché's theorem that the functions $f - w_0$ and $f - w$ have the same number of zeros in the interior of K. The function $f - w_0$ has a zero of order m at z_0 and no other zeros in K. Hence, the function $f - w$ has a total of m zeros in the interior of K. Each of its zeros in K is of order 1 because f' is nonzero in K except possibly at z_0. Thus, f assumes the value w at exactly m distinct points in the interior of K, with multiplicity 1 at each of those points, which is the desired conclusion.

X.14. Consequences of the Local Mapping Theorem

(i) *If f is a nonconstant holomorphic function in a domain G, then f maps open subsets of G onto open sets (briefly, f is an open map).*

In fact, the local mapping theorem implies that the image under f of any disk centered at a point z_0 of G contains a disk centered at $f(z_0)$. The preceding statement follows immediately from this. Note that the maximum modulus principle is a corollary.

(ii) *If the holomorphic function f has a nonvanishing derivative at the point z_0, then there is a disk centered at z_0 in which f is univalent.*

This is the case $m = 1$ of the local mapping theorem. This criterion for univalence is strictly local. The function e^z, for example, has a nowhere vanishing derivative, yet it is not globally univalent.

(iii) *A univalent holomorphic function has a nowhere vanishing derivative.*

For, by the local mapping theorem, if the derivative of a holomorphic function vanishes at some point, then, in every neighborhood of that point, the function fails to be univalent.

Exercise X.14.1. Let the holomorphic function f have a zero of order m at the point z_0. Prove that there is an open disk centered at z_0 in which there is a univalent branch of $f^{1/m}$.

X.15. Inverses

The inverse of a univalent holomorphic function is holomorphic.

In fact, let f be a univalent holomorphic function in the domain G, and let g be the inverse function, defined in $f(G)$. Since f is an open map, the set $f(G)$ is open, and, moreover, if H is any open subset of G, then $g^{-1}(H)$ $(= f(H))$ is open. It follows (by a standard criterion) that the function g is continuous. Because also f' is never 0, one can from this point argue just as in elementary calculus to show that g is differentiable, with

$$g'(w) = \frac{1}{f'(g(w))}, \qquad w \in f(G).$$

The details are left for the reader (see Exercise IV.14.1).

Exercise X.15.1. Let f be a univalent holomorphic function in the domain G, and let g be the inverse function. Let Γ be a simple contour contained with its interior in G. Use the residue theorem to derive the formula

$$g(w) = \frac{1}{2\pi i}\int_\Gamma \frac{zf'(z)}{f(z) - w}dz, \qquad w \in f(\text{int } \Gamma).$$

X.16. Conformal Equivalence

Two domains G_1 and G_2 in $\mathbf{C}$ are said to be conformally equivalent if there is a univalent holomorphic function f in G_1 such that $f(G_1) = G_2$. Such a function f enables one to transport function theory in G_1 to function theory in G_2; it can be viewed as a change of coordinates. Because the class of univalent holomorphic functions is closed under composition and under the formation of inverses, one easily sees that conformal equivalence is an equivalence relation in the standard set-theoretic sense (it is reflexive, symmetric and transitive).

Any open disk in $\mathbf{C}$ is obviously conformally equivalent to the unit disk via a linear function. Using linear-fractional transformations one easily sees that any open half-plane is conformally equivalent to the unit disk. The first exercise below contains further examples of domains that are conformally equivalent to the unit disk.

It follows from Liouville's theorem that the entire plane $\mathbf{C}$ is not conformally equivalent to the unit disk.

Exercise X.16.1. Find a univalent holomorphic map from the open unit disk onto each of the following domains: (i) the upper half-plane; (ii) the whole plane minus the nonpositive real axis; (iii) the strip $0 < \operatorname{Im} z < 1$; (iv) the first quadrant; (v) the intersection of the unit disk with the upper half-plane; (vi) the unit disk minus the segment $[0, 1)$.

Exercise X.16.2. A univalent holomorphic map of a domain onto itself is called a conformal automorphism of the domain. Use Schwarz's lemma to prove that the identity function is the only conformal automorphism of the unit disk that fixes the origin and has a positive derivative at the origin.

Exercise X.16.3. Use the result of the preceding exercise to prove that the only conformal automorphisms of the unit disk are the linear fractional transformations of Exercise III.9.4.

Exercise X.16.4. Prove that the only conformal automorphisms of $\mathbf{C}$ are the nonconstant linear functions.

Exercise X.16.5. Prove that the function $f(z) = \dfrac{z}{(1-z)^2}$ (called the Koebe function) is univalent in the open unit disk. Find the image of the unit disk under f.

X.17. The Riemann Mapping Theorem

Any domain in $\mathbf{C}$ that is conformally equivalent to a simply connected domain must itself be simply connected. This can be seen, for example, on the basis of the characterization of simple connectivity given in Section X.4: the domain G in $\mathbf{C}$ is simply connected if and only if every holomorphic function in G has a primitive in G. (The details are relegated to Exercise X.17.1 below.) A remarkable converse was formulated by G. F. B. Riemann in 1851, but not proved in full generality until 1900 by W. F. Osgood: *Every simply connected domain in* $\mathbf{C}$*, except for* $\mathbf{C}$ *itself, is conformally equivalent to the unit disk.*

In Riemann's time, the modern notion of simple connectivity was yet to be developed, so the theorem envisioned by Riemann was in a technical sense (although not in spirit) less broad than what was eventually proved. The proof Riemann proposed, although basically sound, contained a crucial gap, the eventual patching of which involved major advances in the subject of potential theory (the study of harmonic functions). About a dozen years after Osgood's paper appeared, C. Carathéodory and P. Koebe presented approaches to the theorem which rely on the distinctive methodology of complex function theory, rather than on potential theory. Their methods evolved into an extremely elegant proof which, in its final form, is due to L. Fejér and F. Riesz. It is presented below, after a few more preliminaries.

Exercise X.17.1. Prove that a domain in $\mathbf{C}$ is simply connected if it is conformally equivalent to a simply connected domain.

Exercise* X.17.2. (Uniqueness of the Riemann map.) Let f and g be univalent holomorphic functions in the unit disk, D, such that $f(D) = g(D)$, $f(0) = g(0)$, and $\arg f'(0) = \arg g'(0)$. Prove that $f = g$.

Exercise* X.17.3. Assume the truth of the Riemann mapping theorem. Let G be a simply connected domain in $\mathbf{C}$, different from $\mathbf{C}$, let z_0 be a point of G, and let θ_0 be a real number. Prove that there is a univalent holomorphic function f mapping the unit disk onto G such that $f(0) = z_0$ and $\arg f'(0) = \theta_0$. (The function f is unique, by the preceding exercise.)

X.18. An Extremal Property of Riemann Maps

Suppose G is a simply connected domain, other than $\mathbf{C}$, and suppose there exists a univalent holomorphic function f mapping G onto the unit disk (commonly referred to as a Riemann map of G onto the disk). Let $z_0 = f^{-1}(0)$, and suppose g is any holomorphic function mapping G into the unit disk and satisfying $g(z_0) = 0$. The composite function $g \circ f^{-1}$ is then a holomorphic map of the unit disk into itself which vanishes at the origin. By the corollary to Schwarz's lemma given in Exercise VII.17.1, the derivative of $g \circ f^{-1}$ at 0 has absolute value at most 1, in fact less than 1 unless $g \circ f^{-1}$ is a unimodular constant times the identity function. After an application of the chain rule, one sees that this means $|g'(z_0)| \le |f'(z_0)|$, with equality if and only if $g = \lambda f$ with λ a unimodular constant. The Riemann maps of G onto the unit disk vanishing at z_0 thus maximize the absolute value of the derivative at z_0 among the family of all holomorphic maps of G into the unit disk vanishing at z_0.

The preceding discussion, although premised on the existence of Riemann maps, suggests the strategy behind the existence proof to be given. We shall consider the extremum problem of maximizing $|f'(z_0)|$ among the family of univalent holomorphic functions f that map G into the unit disk and vanish at a given point z_0 of G. We shall prove that the extremum problem has solutions, and then, by a bit of sorcery, prove that the solutions are the desired Riemann maps.

The discussion above, by the way, shows that a Riemann map of G onto the unit disk is uniquely determined by the point it maps to the origin and the argument of its derivative at that point. The reader who has worked Exercises X.17.2 and X.17.3 in the preceding section has already been over this ground.

Existence proofs for solutions of extremum problems often rely on some sort of compactness property. The compactness property to be used in the proof of the Riemann mapping theorem is given in the next section.

X.19. Stieltjes-Osgood Theorem

A locally uniformly bounded sequence of holomorphic functions, defined in a domain in $\mathbf{C}$*, has a locally uniformly convergent subsequence.*

Although the term "locally uniformly bounded" has not been previously defined, the reader has no doubt guessed the definition: a family of functions in a domain is said to be locally uniformly bounded if each point in the domain has a neighborhood in which the family is uniformly bounded (equivalently, if the family is uniformly bounded on each compact subset of the domain).

A family of holomorphic functions in a domain is called a normal family if every sequence from the family has either a locally uniformly convergent subsequence or a subsequence that diverges locally uniformly to ∞. The Stieltjes-Osgood theorem can thus be restated: a locally uniformly bounded family of holomorphic functions is a normal family. The notion of normal family is used extensively in complex function theory.

To prove the Stieltjes-Osgood theorem, let G be our domain, and let $\left(f_n\right)_{n=1}^{\infty}$ be our locally uniformly bounded sequence of holomorphic functions in G. It will be enough to show that, for each closed disk in G, our sequence has a subsequence that converges uniformly on that disk. In fact, suppose this is known. Take a sequence $\left(D_k\right)_{k=1}^{\infty}$ of open disks that cover G and whose closures are contained in G. (For example, take the collection of open disks whose radii are rational, whose centers have rational real and imaginary parts, and whose closures are contained in G. The collection is countable and so can be arranged into a sequence.) By our supposition, the sequence $\left(f_n\right)_{n=1}^{\infty}$ has a subsequence $\left(f_{1,n}\right)_{n=1}^{\infty}$ that converges uniformly on D_1. Similarly, the sequence $\left(f_{1,n}\right)_{n=1}^{\infty}$ has a subsequence $\left(f_{2,n}\right)_{n=1}^{\infty}$ that converges uniformly on D_2. Proceeding recursively, we obtain for each k

a sequence $\left(f_{k,n}\right)_{n=1}^{\infty}$, uniformly convergent on D_k, each sequence being a subsequence of the preceding ones and of the original one. The diagonal sequence $\left(f_{n,n}\right)_{n=1}^{\infty}$ then has the required properties: it is a subsequence of $\left(f_n\right)_{n=1}^{\infty}$, and it converges locally uniformly in G. This is an often used "diagonalization" procedure, most likely familiar to the reader.

Now consider one of our disks D_k. Since $\overline{D_k}$ is contained in G, there is an open disk containing $\overline{D_k}$ whose closure is contained in G, and hence on which our sequence is uniformly bounded. To complete the proof it will suffice to show that our sequence has a subsequence that converges locally uniformly in the larger disk. A linear change of variables maps the larger disk onto the unit disk, reducing our task to that of establishing the following assertion: *A uniformly bounded sequence of holomorphic functions in the unit disk has a subsequence that converges locally uniformly in the disk.*

To prove the assertion, let $(f_n)_1^\infty$ be the given sequence, and let f_n have the power series representation $\sum_{k=0}^{\infty} a_{n,k} z^k$. Choose $M > 0$ such that $|f_n(z)| \leq M$ for $|z| < 1$, for all n. It is asserted that $|a_{n,k}| \leq M$ for all n, k. In fact, letting C_r denote the counterclockwise oriented circle with center 0 and radius r $(0 < r < 1)$, we have

$$|a_{n,k}| = \left| \frac{1}{2\pi i} \int_{C_r} \frac{f_n(z)}{z^{k+1}} dz \right| \leq \frac{M}{r^k},$$

from which the desired inequality follows when we let $r \to 1$.

Thus, each sequence $(a_{n,k})_{n=1}^{\infty}$ $(k = 0, 1, \dots)$ is bounded and so has a convergent subsequence. Now, using the diagonalization procedure in the same way as above, we can obtain an increasing sequence $(n_j)_{j=1}^{\infty}$ of the natural numbers such that the sequence $(a_{n_j,k})_{j=1}^{\infty}$ converges for each k. (The details are relegated to Exercise X.19.1 below.)

Let $a_k = \lim\limits_{j\to\infty} a_{n_j,k}$ $(k = 0, 1, \dots)$. We have $|a_k| \leq M$ for all k, so the power series $\sum_{k=0}^{\infty} a_k z^k$ converges in the unit disk, say to the function f. It will be shown that $f_{n_j} \to f$ locally uniformly in the disk, which will complete the proof of the Stieltjes-Osgood theorem.

Fix r in $(0, 1)$. We show that $f_{n_j}(z) \to f(z)$ uniformly for $|z| \leq r$. Fix $\epsilon > 0$. Take a positive integer m, to be appropriately chosen later. For

$|z| \le r$ we have

$$\begin{aligned}|f(z) - f_{n_j}(z)| &= \left|\sum_{k=0}^{\infty}(a_k - a_{n_j,k})z^k\right| \\ &\le \sum_{k=0}^{m} |a_k - a_{n_j,k}| + 2M \sum_{k=m+1}^{\infty} r^k \\ &= \sum_{k=0}^{m} |a_k - a_{n_j,k}| + \frac{2Mr^{m+1}}{1-r}.\end{aligned}$$

We now fix m so that $2Mr^{m+1}/(1-r) < \frac{\epsilon}{2}$, getting, for $|z| \le r$,

$$|f(z) - f_{n_j}(z)| \le \sum_{k=0}^{m} |a_k - a_{n_j,k}| + \frac{\epsilon}{2}.$$

Finally, since $a_{n_j,k} \to a_k$ for each k as $j \to \infty$, there is a j_0 such that $|a_k - a_{n_j,k}| < \epsilon/2(m+1)$ for $k = 0, \dots, m$ whenever $j > j_0$, and hence so that $|f(z) - f_{n_j}(z)| < \epsilon$ for $|z| \le r$ and $j \ge j_0$, the desired conclusion.

Exercise X.19.1. Fill in the details for the second diagonalization procedure used above by proving that, if $(a_{n,k})_{n=1}^{\infty}$, $k = 1, 2, \dots$, are bounded sequences of complex numbers, then there is an increasing sequence $(n_j)_{j=1}^{\infty}$ of natural numbers such that the sequence $(a_{n_j,k})_{j=1}^{\infty}$ converges for each k.

Exercise X.19.2. Prove that the family of holomorphic functions that map the domain G into the right half-plane, $\operatorname{Re} z > 0$, is a normal family.

Exercise X.19.3. Let G be a bounded domain. Prove that the family of holomorphic functions f in G satisfying $\int\int_G |f(x+iy)|\,dx\,dy \le 1$ is a normal family.

Exercise X.19.4. (Vitali's theorem) Let $\left(f_n\right)_{n=1}^{\infty}$ be a locally uniformly bounded sequence of holomorphic functions in a domain G. Assume that the sequence converges at each point of a set which has a limit point in G. Prove that the sequence converges locally uniformly in G.

Exercise X.19.5. Let f be a bounded holomorphic function in the strip $-1 < \operatorname{Re} z < 1$. Assume that the limit $\lim_{y \to +\infty} f(iy) = c$ exists. Prove that, for each r in $(0,1)$, the limit $\lim_{y \to +\infty} f(x+iy) = c$ exists uniformly for $-r \le x \le r$.

X.20. Proof of the Riemann Mapping Theorem

Let G be a simply connected domain, different from $\mathbf{C}$. Our aim is to prove that there exists a univalent holomorphic map of G onto the open unit disk, which we shall denote by D. Fix a point z_0 in G, and let $\mathcal{F}$ be the family of univalent holomorphic maps of G into D that vanish at z_0.

STEP 1. Our first task is to prove that $\mathcal{F}$ is not empty. For this it will suffice to show that there exists a bounded, univalent holomorphic function in G (for the composite of such a function with a suitable linear function will belong to $\mathcal{F}$). Let c be any point in $\mathbf{C}\backslash G$. Because G is simply connected, there is by X.5 a branch l of the function $\log(z-c)$ in G. The function l is univalent, and the image set $l(G)$ is open. Moreover, $l(G)$ is disjoint from its translate by $2\pi i$. (If $l(z_2) = l(z_1) + 2\pi i$, then

$$z_2 - c = \exp\big(l(z_2)\big) = \exp\big(l(z_1)\big) = z_1 - c,$$

a blatant contradiction.) The function $(l - l(z_0) - 2\pi i)^{-1}$ is thus bounded, univalent, and holomorphic in G.

STEP 2. Now that we know $\mathcal{F}$ is not empty, it makes sense to consider the problem of maximizing $|f'(z_0)|$ over the functions f in $\mathcal{F}$. Let

$$\rho = \sup\{|f'(z_0)| : f \in \mathcal{F}\}.$$

The number ρ is positive because the functions in $\mathcal{F}$ have nonvanishing derivatives. (The finiteness of ρ will be apparent momentarily.) Let $\big(f_n\big)_{n=1}^{\infty}$ be a sequence in $\mathcal{F}$ such that $|f_n'(z_0)| \to \rho$. By the Stieltjes-Osgood theorem, this sequence has a locally uniformly convergent subsequence. So as not to complicate our notation needlessly, we may assume the original sequence converges locally uniformly. Let f denote the limit function. By the Weierstrass convergence theorem VII.15, f is holomorphic, and $f_n' \to f'$ locally uniformly. Thus $|f'(z_0)| = \rho$, which implies in particular that the function f is nonconstant. Since the functions f_n are univalent, it now follows by Hurwitz's theorem (Exercise X.12.5) that f is univalent. The function f is bounded in modulus by 1, being the limit of such functions, and it vanishes at z_0. By the maximum modulus principle, f maps G into D. The function f thus belongs to the family $\mathcal{F}$; it is a solution of our extremal problem.

STEP 3. Finally, we complete the proof of the Riemann mapping theorem by showing that $f(G) = D$. The argument involves some sleight of hand devised by Koebe (his "Schmiegungsverfahren").

We argue by contradiction, supposing $f(G)$ is a proper subset of D. Let a be any point in $D\backslash f(G)$, and define the function h_1 by

$$h_1 = \frac{f-a}{1-\bar{a}f}.$$

The function h_1 is a univalent holomorphic map of G into D, being the composite of f with a conformal automorphism of D. Also h_1 does not vanish in G, so, because G is simply connected, there is a branch h_2 of the function $\sqrt{h_1}$ in G. It also is a univalent holomorphic map of G into D.

Lastly, we create a function g in our family $\mathcal{F}$ by setting $b = h_2(z_0)$ and defining

$$g = \frac{h_2 - b}{1 - \bar{b}h_2}.$$

Now, after a bit of algebra (relegated to Exercise X.20.1 below) one finds that

$$f = \left(\frac{g + \alpha}{1 + \bar{\alpha}g}\right)g,$$

where $\alpha = \dfrac{2b}{1 + |b|^2}$. This gives $f'(z_0) = \alpha g'(z_0)$, so $\rho = |f'(z_0)| < |g'(z_0)|$ since $|\alpha| < 1$, a contradiction. This concludes the proof of the Riemann mapping theorem.

The proof just given, while undeniably elegant, is at the same time highly nonconstructive. It does not address the question of how to find usable approximations to Riemann maps in concrete situations. That question is an active topic in numerical analysis.

The Riemann mapping theorem is the starting point of the subject of geometric function theory, a vast endeavor, very vigorous today despite its relatively advanced age. The reader can find more information in the books on complex analysis listed among the references.

Exercise X.20.1. Perform the algebra needed in the preceding proof to obtain the expression for the function f in terms of the function g.

Exercise X.20.2. Assume, retaining the notations of the preceding proof, that $\left(f_n\right)_{n=1}^{\infty}$ is a sequence in $\mathcal{F}$ such that $|f_n'(z_0)| \to \rho$ and $\lim\limits_{n\to\infty} f_n'(z_0)$ exists. Prove that the sequence $\left(f_n\right)_{n=1}^{\infty}$ converges locally uniformly in G.

Exercise X.20.3. Let $\left(G_n\right)_{n=1}^{\infty}$ be an increasing sequence of simply connected domains whose union, G, is not the whole plane. Let z_0 be a point of G_1. For each n, let f_n be the Riemann map of D onto G_n satisfying $f_n(0) = z_0$ and $\arg f_n'(0) = 0$. Prove that the sequence $\left(f_n\right)_{n=1}^{\infty}$ converges locally uniformly in D to the Riemann map f of D onto G satisfying $f(0) = z_0$ and $\arg f'(0) = 0$.

X.21. Simple Connectivity Again

If G is a simply connected domain in $\mathbf{C}$, then every closed curve in G is null homotopic in G.

With this we see finally that the definition we used for simple connectivity, in the context of planar domains, is equivalent to the definition using homotopy. The equivalence is nontrivial; we will use the Riemann mapping theorem to establish the implication above. Namely, a simple argument (left to the reader) establishes the implication for the cases $G = \mathbf{C}$ and $G = D$. If $G \neq \mathbf{C}$ then G is conformally equivalent to D, hence homeomorphic to D. Since a homeomorphism preserves homotopy, the validity of the implication for D implies its validity for G.

Exercise X.21.1. Prove that a domain G in $\mathbf{C}$ is simply connected if and only if its extended boundary (the boundary relative to $\overline{\mathbf{C}}$) is connected.

Appendix 1. Sufficient condition for differentiability

Let the real-valued function u, defined in an open set containing the point z_0 of $\mathbf{C}$, have first partial derivatives that are continuous at z_0. Then u is differentiable at z_0.

The proof relies on the mean value theorem of calculus: *If the real-valued function f is differentiable on the interval $[a, b]$, then there is a point c in (a, b) such that $f(b) - f(a) = f'(c)(b - a)$.*

We write $z_0 = x_0 + i\,y_0$. To avoid possible notational confusion, we let u_1 and u_2 denote the partial derivatives of u in the real and imaginary directions, respectively ($u_1 = \frac{\partial u}{\partial x}$, $u_2 = \frac{\partial u}{\partial y}$). Using the mean value theorem we obtain, for $z = x + iy$ near z_0,

$$\begin{aligned} u(z) - u(z_0) &= u(x, y) - u(x_0, y_0) \\ &= u(x, y) - u(x, y_0) + u(x, y_0) - u(x_0, y_0) \\ &= u_2(x, \eta_z)(y - y_0) + u_1(\xi_z, y_0)(x - x_0), \end{aligned}$$

where η_z is between y_0 and y, and ξ_z is between x_0 and x. Thus,

$$\begin{aligned} &u(x, y) - u(x_0, y_0) - u_1(x_0, y_0)(x - x_0) - u_2(x_0, y_0)(y - y_0) \\ &= [u_2(x, \eta_z) - u_2(x_0, y_0)](y - y_0) + [u_1(\xi_z, y_0) - u_1(x_0, y_0)](x - x_0), \end{aligned}$$

from which it follows that

$$\frac{|u(x,y) - u(x_0,y_0) - u_1(x_0,y_0)(x - x_0) - u_2(x_0,y_0)(y - y_0)|}{|z - z_0|}$$
$$\leq |u_2(x,\eta_z) - u_2(x_0,y_0)| + |u_1(\xi_z,y_0) - u_1(x_0,y_0)|.$$

As $z \to z_0$ we have $(x,\eta_z) \to (x_0,y_0)$ and $(\xi_z,y_0) \to (x_0,y_0)$, so by the continuity of u_1 and u_2 at z_0 the quantity on the right side in the preceding inequality tends to 0, which gives the desired conclusion.

Appendix 2. Two instances of the chain rule

1. *Let the function f be holomorphic in the open set G. Let $\gamma : I \to G$ be a curve, differentiable at the point t_0 of I. Then the composite curve $f \circ \gamma$ is differentiable at t_0, with $(f \circ \gamma)'(t_0) = f'(\gamma(t_0))\gamma'(t_0)$.*

We let $z_0 = \gamma(t_0)$. By the differentiability of f at z_0 we have

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + R_1(z),$$

where $R_1(z)/(z - z_0) \to 0$ as $z \to z_0$. By the differentiability of γ at t_0 we have

$$\gamma(t) = \gamma(t_0) + \gamma'(t_0)(t - t_0) + R_2(t),$$

where $R_2(t)/(t - t_0) \to 0$ as $t \to t_0$. Thus

$$f(\gamma(t)) = f(\gamma(t_0)) + f'(z_0)(\gamma(t) - \gamma(t_0)) + R_1(\gamma(t))$$

$$= f(\gamma(t_0)) + f'(z_0)\gamma'(t_0)(t - t_0) + R_3(t),$$

where

$$R_3(t) = f'(z_0)R_2(t) + R_1(\gamma(t)).$$

We need to show $R_3(t)/(t - t_0) \to 0$ as $t \to t_0$. We have

$$\frac{R_3(t)}{t - t_0} = f'(z_0)\left(\frac{R_2(t)}{t - t_0}\right) + \frac{R_1(\gamma(t))}{t - t_0}.$$

The first summand on the right clearly tends to 0 as $t \to t_0$. As for the second summand, it equals 0 if $\gamma(t) = z_0$, and otherwise it can be written as

$$\frac{R_1(\gamma(t))}{\gamma(t) - \gamma(t_0)} \left(\frac{\gamma(t) - \gamma(t_0)}{t - t_0} \right),$$

from which one sees that it also tends to 0 as $t \to t_0$.

2. *Let $f = u + iv$ be a C^1 function defined in an open subset G of* $\mathbf{C}$. *Let γ be a differentiable curve in G, t_0 a point where γ is defined, $z_0 = \gamma(t_0)$, $a = \frac{\partial f(z_0)}{\partial z}$, $b = \frac{\partial f(z_0)}{\partial \bar{z}}$.* Then

$$(f \circ \gamma)'(t_0) = a\gamma'(t_0) + b\,\overline{\gamma'(t_0)}.$$

We write $\gamma = \alpha + i\,\beta$. Below, for the sake of conciseness in notation, dependencies on the parameter t and the variables x and y are suppressed. By the chain rule of multivariable calculus,

$$\frac{\mathrm{d}u(\gamma)}{\mathrm{d}t} = \frac{\mathrm{d}u(\alpha,\beta)}{\mathrm{d}t} = \frac{\partial u}{\partial x}\alpha' + \frac{\partial u}{\partial y}\beta'$$
$$\frac{\mathrm{d}v(\gamma)}{\mathrm{d}t} = \frac{\mathrm{d}v(\alpha,\beta)}{\mathrm{d}t} = \frac{\partial v}{\partial x}\alpha' + \frac{\partial v}{\partial y}\beta'.$$

Hence,

$$\begin{aligned}(f\circ\gamma)' &= \frac{\partial f}{\partial x}\alpha' + \frac{\partial f}{\partial y}\beta' \\ &= \left(\frac{\partial f}{\partial z} + \frac{\partial f}{\partial \bar{z}}\right)\alpha' + i\left(\frac{\partial f}{\partial z} - \frac{\partial f}{\partial \bar{z}}\right)\beta' \\ &= \frac{\partial f}{\partial z}(\alpha' + i\beta') + \frac{\partial f}{\partial \bar{z}}(\alpha' - i\beta') \\ &= \frac{\partial f}{\partial z}\gamma' + \frac{\partial f}{\partial \bar{z}}\overline{\gamma'}.\end{aligned}$$

This holds in particular for $t = t_0$, which is the desired conclusion.

Appendix 3. Groups, and linear-fractional transformations

This appendix is aimed at readers unfamiliar with the notion of a group, in the hope that it will help them understand some of the statements in Chapter III.

A group is by definition a set G equipped with a binary operation

$$G \times G \to G$$
$$(a, b) \to ab$$

(the group operation) having the properties: (i) $a(bc) = (ab)c$ for all a, b, c; (ii) There is an element e in G (the identity element) such that $ea = ae = a$ for all a; (iii) Each element a has an inverse a^{-1}, an element of G satisfying $a^{-1}a = aa^{-1} = e$. As shown in Section III.2, the preceding axioms are satisfied by the set of linear-fractional transformations equipped with the binary operation of composition, the identity element in this case being the identity transformation. The axioms are also satisfied by $GL(2, \mathbf{C})$, the set of nonsingular two-by-two complex matrices equipped with the operation of matrix multiplication; the identity here is the identity matrix.

An isomorphism of a group G_1 onto a group G_2 is a one-to-one map J of G_1 onto G_2 that preserves the group operation: $J(ab) = J(a)J(b)$ for all a, b in G_1. If such a map J exists then G_1 and G_2 are said to be isomorphic. The map of $GL(2, \mathbf{C})$ onto the group of linear-fractional transformations that sends each matrix to its corresponding transformation preserves the group operation but, while surjective, it is not one-to-one, so it is not an isomorphism; it is what is called a homomorphism. The quotient group $GL(2, \mathbf{C})/(\mathbf{C}\backslash\{0\})I_2$ mentioned in Section III.2 has as its elements equivalence classes of $GL(2, \mathbf{C})$, where two matrices are regarded as equivalent if they are scalar multiples of each other (i.e., if they induce the same linear-fractional transformation). The group operation on the quotient group is defined in the obvious way, and the statement that the quotient group is isomorphic to the group of linear-fractional transformations is simple to verify. This isomorphism was mentioned in Chapter 3 by way of orientation for readers familiar with group theory; it is not used subsequently.

Section III.5 concerns the property that any three distinct points z_1, z_2, z_3 can be mapped in order to any other three distinct points w_1, w_2, w_3 by a linear-fractional transformation. By way of a proof, the property was established for a particular choice of w_1, w_2, w_3 (namely ∞, 0, 1), after the statement that this would be sufficient. To justify that statement, consider arbitrary z_1, z_2, z_3 and arbitrary w_1, w_2, w_3. According to the special case treated, there is a linear-fractional transformation ϕ_1 mapping z_1, z_2, z_3 to ∞, 0, 1, and there is a linear-fractional transformation ϕ_2 mapping w_1, w_2, w_3 to ∞, 0, 1. The linear-fractional transformation $\phi_2^{-1} \circ \phi_1$ then maps z_1, z_2, z_3 to w_1, w_2, w_3.

Appendix 4. Differentiation under the integral sign

Let the function $F : [r_1, r_2] \times [t_1, t_2] \to \mathrm{R}$ *be continuous and have a continuous first partial derivative with respect to its first variable. Let the function* $I : [r_1, r_2] \to \mathrm{R}$ *be defined by*

$$I(r) = \int_{t_1}^{t_2} F(r, t)dt. \tag{1}$$

Then I *is differentiable, and*

$$\frac{dI(r)}{dr} = \int_{t_1}^{t_2} \frac{\partial F(r,t)}{\partial r} dt.$$

This result was needed in Section VIII.4, in the proof of Cauchy's theorem for a pair of concentric circles. In proving (1), we simplify the notation by letting F_1 denote the first partial derivative of F with respect to the first variable. Fix r_0 in $[r_1, r_2]$. We have

$$\begin{aligned} &\frac{I(r) - I(r_0)}{r - r_0} - \int_{t_1}^{t_2} F_1(r_0, t)dt \\ &= \int_{t_1}^{t_2} \left[\frac{F(r,t) - F(r_0, t)}{r - r_0} - F_1(r_0, t) \right] dt. \end{aligned} \tag{2}$$

As $r \to r_0$, the integrand in the integral on the right tends pointwise to 0. To obtain the desired result we need to prove that the integral tends to 0, for which we need to estimate the rate at which the integrand tends to 0.

Fix $\epsilon > 0$. The function F_1, being continuous on the compact set $[r_1, r_2] \times [t_1, t_2]$, is uniformly continuous there. Hence, there is a $\delta > 0$ such that $|F_1(r,t) - F_1(r',t')| < \epsilon$ whenever the distance between (r,t) and (r',t') is less than δ. Suppose $|r - r_0| < \delta$. For each t, there is by the mean value theorem a number r_t between r_0 and r such that

$$\frac{F(r,t) - F(r_0,t)}{r - r_0} = F_1(r_t, t).$$

By the choice of δ, this implies that

$$\left| \frac{F(r,t) - F(r_0,t)}{r - r_0} - F_1(r_0,t) \right| < \epsilon.$$

So, whenever $|r - r_0| < \delta$ the integrand in the integral on the right hand side of (2) has absolute value less than ϵ for all t, implying that the integral has absolute value less than $\epsilon(t_2 - t_1)$. As ϵ is arbitrary, the desired conclusion follows.

References

1. Ahlfors, L. V., *Complex Analysis*, McGraw-Hill, 1979.
2. Axler, S., Bourdon, P., and Ramey, W., *Harmonic Function Theory*, Springer-Verlag, 1992.
3. Burckel, R. B., *An Introduction to Classical Complex Analysis, Vol. 1*, Birkhäuser, 1979.
4. Conway, J. B., *Functions of One Complex Variable*, Springer-Verlag, 1978.
5. Marsden, J. E., and Hoffman, M. J., *Basic Complex Analysis*, W. H. Freeman, 1987.
6. Newman, M. H. A., *Elements of the Topology of Plane Sets of Points*, Cambridge, 1961.
7. Rudin, W., *Principles of Mathematical Analysis*, McGraw-Hill, 1976.
8. Saks, S., and Zygmund, A., *Analytic Functions*, Elsevier, 1971.

Index

Complex Function Theory is a concise and rigorous introduction to the theory of functions of a complex variable. Written in a classical style, it is in the spirit of the books by Ahlfors and by Saks and Zygmund. Being designed for a one-semester course, it is much shorter than many of the standard texts. Sarason covers the basic material through Cauchy's theorem and applications, plus the Riemann mapping theorem. It is suitable for either an introductory graduate course or an undergraduate course for students with adequate preparation.

The first edition was published with the title Notes on Complex Function Theory.

ISBN 978-0-8218-4428-1

9 780821 844281

MBK/49