S. Ponnusamy
Herb Silverman

Complex Variables with Applications

Birkhäuser
Boston • Basel • Berlin

S. Ponnusamy
Indian Institute of Technology, Madras
Department of Mathematics
Chennai, 600 036
India

Herb Silverman
College of Charleston
Department of Mathematics
Charleston, SC 29424
U.S.A.

Cover design by Alex Gerasev.

Mathematics Subject Classification (2000): 11A06, 11M41, 30-XX, 32-XX (primary); 26Axx, 40Axx, 26Bxx, 33Bxx, 26Cxx, 28Cxx, 31Axx, 35Axx, 37F10, 45E05, 76M40 (secondary)

Library of Congress Control Number: 2006927602

ISBN-10: 0-8176-4457-1 eISBN: 0-8176-4513-6
ISBN-13: 978-0-8176-4457-4

Printed on acid-free paper.

Printed in the United States of America. (TXQ/MP)

9 8 7 6 5 4 3 2 1

www.birkhauser.com

To my father, Saminathan Pillai

—S. Ponnusamy

To my wife, Sharon Fratepietro

—Herb Silverman

Preface

The student, who seems to be engulfed in our culture of specialization, too quickly feels the necessity to establish an "area" of special interest. In keeping with this spirit, academic bureaucracy has often forced us into a compartmentalization of courses, which pretend that linear algebra is disjoint from modern algebra, that probability and statistics can easily be separated, and even that advanced calculus does not build from elementary calculus.

This book is written from the point of view that there is an interdependence between real and complex variables that should be explored at every opportunity. Sometimes we will discuss a concept in real variables and then generalize to one in complex variables. Other times we will begin with a problem in complex variables and reduce it to one in real variables. Both methods—generalization and specialization—are worthy of careful consideration.

We expect "complex" numbers to be difficult to comprehend and "imaginary" units to be shrouded in mystery. Hopefully, by staying close to the real field, we shall overcome this regrettable terminology that has been thrust upon us. The authors wish to create a spiraling effect that will first enable the reader to draw from his or her knowledge of advanced calculus in order to demystify complex variables, and then use this newly acquired understanding of complex variables to master some of the elements of advanced calculus.

We will also compare, whenever possible, the analytic and geometric character of a concept. This naturally leads us to a discussion of "rigor". The current trend seems to be that anything analytic is rigorous and anything geometric is not. This dichotomy moves some authors to strive for "rigor" at the expense of rich geometric meaning, and other authors to endeavor to be "intuitive" by discussing a concept geometrically without shedding any analytic light on it. Rigor, as the authors see it, is useful only insofar as it clarifies rather than confounds. For this reason, geometry will be utilized to illustrate analytic concepts, and analysis will be employed to unravel geometric notions, without regard to which approach is the more rigorous.

Sometimes, in an attempt to motivate, a discussion precedes a theorem. Sometimes, in an attempt to illuminate, remarks about key steps and possible implications follow a theorem. No apologies are made for this lack of terseness surrounding difficult theorems. While brevity may be the soul of wit, it is not the soul of insight into delicate mathematical concepts. In recognition of the primary importance of observing relationships between different approaches, some theorems are proved in several different ways. In this book, traveling quickly to the frontiers of mathematical knowledge plays a secondary role to the careful examination of the road taken and alternative routes that lead to the same destination.

A word should be said about the questions at the end of each section. The authors feel deeply that mathematics should be questioned—not only for its internal logic and consistency, but for the reasons we are led where we are. Does the conclusion seem "reasonable"? Did we expect it? Did the steps seem natural or artificial? Can we re-prove the result a different way? Can we state intuitively what we have proved? Can we draw a picture?[1]

"Questions", as used at the end of each section, cannot easily be categorized. Some questions are simple and some are quite challenging; some are specific and some are vague; some have one possible answer and some have many; some are concerned with what has been proved and some foreshadow what will be proved. Do all these questions have anything in common? Yes. They are all meant to help the student think, understand, create, and question. It is hoped that the questions will also be helpful to the teacher, who may want to incorporate some of them into his or her lectures.

Less need be said about the exercises at the end of each section because exercises have always received more favorable publicity than have questions. Very often the difference between a question and an exercise is a matter of terminology. The abundance of exercises should help to give the student a good indication of how well the material in the section has been understood.

The prerequisite is at least a shaky knowledge of advanced calculus. The first nine chapters present a solid foundation for an introduction to complex variables. The last four chapters go into more advanced topics in some detail, in order to provide the groundwork necessary for students who wish to pursue further the general theory of complex analysis.

If this book is to be used as a one-semester course, Chapters 5, 6, 7, 8, and 9 should constitute the core. Chapter 1 can be covered rapidly, and the concepts in Chapter 2 need be introduced only when applicable in latter chapters. Chapter 3 may be omitted entirely, and the mapping properties in Chapter 4 may be omitted.

We wanted to write a mathematics book that omitted the word "trivial". Unfortunately, the Riemann hypothesis, stated on the last page of the text,

[1] For an excellent little book elaborating on the relationship between questioning and creative thinking, see G. Polya, *How to Solve It,* second edition, Princeton University press, Princeton, New Jersey, 1957.

could not have been mentioned without invoking the standard terminology dealing with the *trivial zeros* of the Riemann zeta function. But the spirit, if not the letter, of this desire has been fulfilled. Detailed explanations, remarks, worked-out examples and insights are plentiful. The teacher should be able to leave sections for the student to read on his/her own; in fact, this book might serve as a self-study text.

A teacher's manual containing more detailed hints and solutions to questions and exercises is available. The interested teacher may contact us by e-mail and receive a pdf version.

We wish to express our thanks to the Center for Continuing Education at the Indian Institute of Technology Madras, India, for its support in the preparation of the manuscript.

Finally, we thank Ann Kostant, Executive Editor, Birkhäuser, who has been most helpful to the authors through her quick and efficient responses throughout the preparation of this manuscript.

S. Ponnusamy
IIT Madras, India

Herb Silverman
College of Charleston, USA

June 2005

Contents

1

Algebraic and Geometric Preliminaries

The mathematician Euler once said, "God made integers, all else is the work of man." In this chapter, we have advanced in the evolutionary process to the real number system. We partially characterize the real numbers and then, alas, find an imperfection. The quadratic equation $x^2+1=0$ has no solution.

A new day arrives, the complex number system is born. We view a complex number in several ways: as an element in a field, as a point in the plane, and as a two-dimensional vector. Each way is useful and in each way we see an unmistakable resemblance of the complex number system to its parent, the real number system. The child seems superior to its parent in every way except one—it has no order. This sobering realization creates a new respect for the almost discarded parent.

The moral of this chapter is clear. As long as the child follows certain guidelines set down by its parent, it can move in new directions and teach us many things that the parent never knew.

1.1 The Complex Field

We begin our study by giving a very brief motivation for the origin of complex numbers. If all we knew were positive integers, then we could not solve the equation $x+2=1$. The introduction of negative integers enables us to obtain a solution. However, knowledge of every integer is not sufficient for solving the equation $2x-1=2$. A solution to this equation requires the study of *rational numbers.*

While all linear equations with integers coefficients have rational solutions, there are some quadratics that do not. For instance, *irrational numbers* are needed to solve $x^2-2=0$. Going one step further, we can find quadratic equations that have no real (rational or irrational) solutions. The equation $x^2+1=0$ has no real solutions because the square of any real number is nonnegative. In order to solve this equation, we must "invent" a number whose square is -1. This number, which we shall denote by $i = \sqrt{-1}$, is called an *imaginary unit.*

Our sense of logic rebels against just "making up" a number that solves a particular equation. In order to place this whole discussion in a more rigorous setting, we will define operations involving combinations of real numbers and imaginary units. These operations will be shown to conform, as much as possible, to the usual rules for the addition and multiplication of real numbers. We may express any ordered pair of real numbers (a, b) as the "complex number"

$$a + bi \quad \text{or} \quad a + ib. \tag{1.1}$$

The set of complex numbers is thus defined as the set of all ordered pairs of real numbers. The notion of equality and the operations of addition and multiplication are defined as follows:[1]

$$\begin{aligned} (a_1, b_1) = (a_2, b_2) &\Longleftrightarrow a_1 = a_2,\ b_1 = b_2, \\ (a_1, b_1) + (a_2, b_2) &= (a_1 + a_2, b_1 + b_2), \\ (a_1, b_1)(a_2, b_2) &= (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1). \end{aligned}$$

The definition for the multiplication is more natural than it appears to be, for if we denote the complex numbers of the form (1.1), multiply as we would real numbers, and use the relation $i^2 = -1$, we obtain

$$(a_1 + ib_1)(a_2 + ib_2) = a_1 a_2 - b_1 b_2 + i(a_1 b_2 + a_2 b_1).$$

Several observations should be made at this point. First, note that the formal operations for addition and multiplication of complex numbers do not depend on an imaginary number i. For instance, the relation $i^2 = -1$ can be expressed as $(0, 1)(0, 1) = (-1, 0)$. The symbol i has been introduced purely as a matter of notational convenience. Also, note that the order pair $(a, 0)$ represents the real number a, and that the relations

$$(a, 0) + (b, 0) = (a + b, 0) \quad \text{and} \quad (a, 0)(b, 0) = (ab, 0)$$

are, respectively, addition and multiplication of real numbers. Some of the essential properties of real numbers are as follows: Both the *sum* and *product* of real numbers are real numbers, and the *order* in which either operation is performed may be *reversed.* That is, for real numbers a and b, we have the commutative laws

$$a + b = b + a \quad \text{and} \quad a \cdot b = b \cdot a. \tag{1.2}$$

The associative laws

$$a + (b + c) = (a + b) + c \quad \text{and} \quad a \cdot (b \cdot c) = (a \cdot b) \cdot c, \tag{1.3}$$

[1] The symbol $\Longleftrightarrow$ stands for "if and only if" or "equivalent to."

and the distributive law

$$a \cdot (b + c) = a \cdot b + a \cdot c \tag{1.4}$$

also holds for all real numbers a, b, and c. The numbers 0 and 1 are, respectively, the *additive* and *multiplicative identities*. The *additive inverse* of a is $-a$, and the *multiplicative inverse* of $a\,(\neq 0)$ is the real number $a^{-1} = 1/a$. Stated more concisely, the real numbers form a *field* under the operations of addition and multiplication.

Of course, the real numbers are not the only system that forms a field. The rational numbers are easily seen to satisfy the above conditions for a field. What is important in this chapter is that the complex numbers also form a field. The additive identity is $(0, 0)$, and the additive inverse of (a, b) is $(-a, -b)$. The multiplicative inverse of $(a, b) \neq (0, 0)$ is

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

We leave the confirmation that the complex numbers satisfy all the axioms for a field as an exercise for the reader.

The discerning math student should not be satisfied with the mere verification of a proof. He/she should also have a "feeling" as to why the proof works. Did the reader ask why the multiplicative inverse of (a, b) might be expected to be

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

Let us go through a possible line of reasoning. If we write the inverse of $(a, b) = a + bi$ as

$$(a + ib)^{-1} = \frac{1}{a + ib},$$

then we want to find a complex number $c + di$ such that

$$\frac{1}{a + ib} = c + id.$$

By cross multiplying, we obtain $ac + i^2bd + i(ad + bc) = 1$, or

$$\begin{cases} ac - bd = 1, \\ ad + bc = 0. \end{cases}$$

The solution to these simultaneous equation is

$$c = \frac{a}{a^2 + b^2}, \quad d = -\frac{b}{a^2 + b^2}.$$

Can the reader think of other reasons to suspect that

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

Let $z = (x, y)$ be a complex number. Then x and y are called the *real part* of z, $\operatorname{Re} z$, and the *imaginary part* of z, $\operatorname{Im} z$, respectively. Denote the set of real numbers by $\mathbb{R}$ and the set of complex numbers by $\mathbb{C}$. There is a one-to-one correspondence between $\mathbb{R}$ and a subset of $\mathbb{C}$, represented by $x \leftrightarrow (x, 0)$ for $x \in \mathbb{R}$, which preserves the operations of addition and multiplication. Hence we will use the real number x and the ordered pair $(x, 0)$ interchangeably. We will also denote the ordered pair $(0, 1)$ by i. Because a complex number is an ordered pair of real numbers, we use the terms $\mathbb{C} = \mathbb{R}^2$ or $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ interchangeably. Thus $\mathbb{R} \times 0$ is a subset of $\mathbb{C}$ consisting of the real numbers.

As noted earlier, an advantage of the field $\mathbb{C}$ is that it contains a root of $z^2 + 1 = 0$. In Chapter 8 we will show that any polynomial equation $a_0 + a_1 z + \cdots + a_n z^n = 0$ has a solution in $\mathbb{C}$. But this extension from $\mathbb{R}$ to $\mathbb{C}$ is not without drawbacks. There is an important property of the real field that the complex field lacks. If $a \in \mathbb{R}$, then exactly *one* of the following is true:

$$a = 0, \quad a > 0, \quad -a > 0 \quad \text{(trichotomy)}.$$

Furthermore, the sum and the product of two positive real numbers is positive (closure).

A field with an order relation $<$ that satisfies the trichotomy law and these two additional conditions is said to be *ordered.* In an ordered field, like the real or rational numbers, we are furnished with a natural way to compare any two elements a and b. Either a is less than b ($a < b$), or a is equal to b ($a = b$), or a is greater than b ($a > b$). Unfortunately, no such relation can be imposed on the complex numbers, for suppose the complex numbers are ordered; then either i or $-i$ is positive. According to the closure rule, $i^2 = (-i)^2 = -1$ is also positive. But 1 must be negative if -1 is positive. However, this violates the closure rule because $(-1)^2 = 1$.

To sum up, there is a complex field that contains a real field that contains a rational field. There are advantages and disadvantages to studying each field. It is not our purpose here to state properties that uniquely determine each field, although this most certainly can be done.

Questions 1.1.

1. Can a field be finite?
2. Can an ordered field be finite?
3. Are there fields that properly contain the rationals and are properly contained in the reals?
4. When are two complex numbers z_1 and z_2 equal?
5. What complex numbers may be added to or multiplied by the complex number $a + ib$ to obtain a real number?
6. How can we separate the quotient of two complex numbers into its real and imaginary parts?

7. What can we say about the real part of the sum of the two complex numbers? What about the product?
8. What kind of implications are there in defining a complex number as an ordered pair?
9. If a polynomial of degree n has at least one solution, can we say more?
10. If we try to define an ordering of the complex numbers by saying that $(a, b) > (c, d)$ if $a > b$ and $c > d$, what order properties are violated?
11. Can any ordered field have a solution to $x^2 + 1 = 0$?

Exercises 1.2.

1. Show that the set of real numbers of the form $a + b\sqrt{2}$, where a and b are rational, is an ordered field.
2. If a and b are elements in a field, show that $ab = 0$ if and only if either $a = 0$ or $b = 0$.
3. Suppose a and b are elements in an ordered field, with $a < b$. Show that there are infinitely many elements between a and b.
4. Find the values of

   (a) $(-2, 3)(4, -1)$ (b) $(1 + 2i)\{3(2 + i) - 2(3 + 6i)\}$
   (c) $(1 + i)^3$ (d) $(1 + i)^4$
   (e) $(1 + i)^n - (1 - i)^n$.
5. Express the following in the form $x + iy$:

   (a) $(1 + i)^{-5}$ (b) $(3 - 2i)/(1 - i)$
   (c) $e^{i\pi/2} + \sqrt{2}e^{i\pi/4}$ (d) $(1 + i)e^{i\pi/6}$
   (e) $\dfrac{a + ib}{a - ib} - \dfrac{a - ib}{a + ib}$ (f) $\dfrac{3 + 5i}{7 + i} + \dfrac{1 + i}{4 + 3i}$
   (g) $(2 + i)^2 + (2 - i)^2$ (h) $\dfrac{(4 + 3i)\sqrt{3 + 4i}}{3 + i}$
   (i) $\dfrac{(ai^{40} - i^{17})}{-1 + i}$, ($a$–real) (j) $(-1 + i\sqrt{3})^{60}$
   (k) $\dfrac{\sqrt{1 + a^2} + ia}{a - i\sqrt{1 + a^2}}$, ($a$–real) (l) $\dfrac{(\sqrt{3} - i)^2(1 + i)^5}{(\sqrt{3} + i)^4}$.
6. Show that
$$\left(\frac{-1 \pm \sqrt{3}}{2}\right)^3 = 1 \quad \text{and} \quad \left(\frac{\pm 1 \pm i\sqrt{3}}{2}\right)^6 = 1$$
   for all combinations of signs.
7. For any integers k and n, show that $i^n = i^{n+4k}$. How many distinct values can be assumed by i^n?

1.2 Rectangular Representation

Just as a real number x may be represented by a point on a line, so may a complex number $z = (x, y)$ be represented by a point in the plane (Figure 1.1).

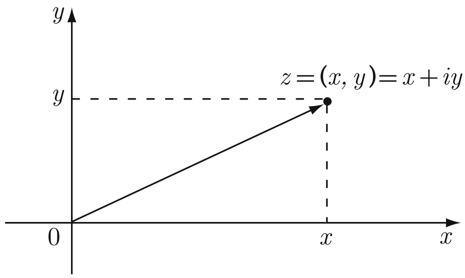


Figure 1.1. Cartesian representation of z in plane

Each complex number corresponds to one and only one point. Thus the terms complex number and point in the plane are used interchangeably. The x and y axes are referred to as the *real axis* and the *imaginary axis*, while the xy plane is called the *complex plane* or the z plane.

There is yet another interpretation of the complex numbers. Each point (x, y) of the complex plane determines a two-dimensional *vector* (directed line segment) from $(0, 0)$, the initial point, to (x, y), the terminal point. Thus the complex number may be represented by a vector. This seems natural in that the definition chosen for addition of complex numbers corresponds to vector addition; that is,

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

Geometrically, vector addition follows the so-called *parallelogram rule*, which we illustrate in Figure 1.2. From the point z_1, construct a vector equal in magnitude and direction to the vector z_2. The terminal point is the vector $z_1 + z_2$. Alternatively, if a vector equal in magnitude and direction to z_1 is joined to the vector z_2, the same terminal point is reached. This illustrates the *commutative property* of vector addition. Note that the vector $z_1 + z_2$ is a diagonal of the parallelogram formed. What would the other diagonal represent?

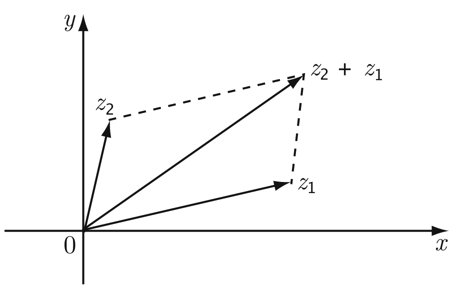


Figure 1.2. Illustration for parallelogram law

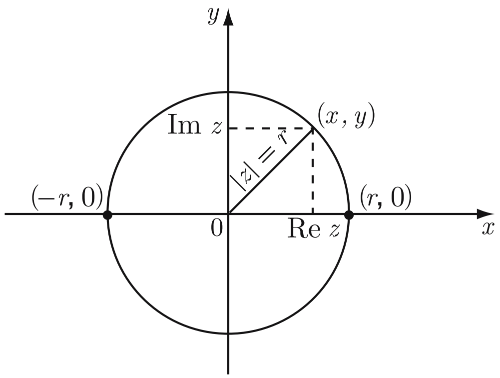


Figure 1.3. Modulus of a complex number z

By the *magnitude* (length) of the vector (x, y) we mean the distance of the point $z = (x, y)$ from the origin. This distance is called the *modulus* or *absolute value* of the complex number z, and denoted by $|z|$; its value is $\sqrt{x^2 + y^2}$. For each positive real number r, there are infinitely many distinct values (x, y) whose absolute value is $r = |z|$, namely the points on the circle $x^2 + y^2 = r^2$. Two of these points, $(r, 0)$ and $(-r, 0)$, are real numbers so that this definition agrees with the definition for the absolute value in the real field (see Figure 1.3).

Note that, for $z = (x, y)$,

$$\begin{cases} |x| = |\mathrm{Re}\, z| \leq |z|, \\ |y| = |\mathrm{Im}\, z| \leq |z|. \end{cases}$$

The distance between any two points $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ is

$$|z_2 - z_1| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

The triangle inequalities

$$\begin{cases} |z_1 + z_2| \leq |z_1| + |z_2|, \\ |z_1 - z_2| \geq \big|\,|z_1| - |z_2|\,\big| \end{cases}$$

say, geometrically, that no side of a triangle is greater in length than the sum of the lengths of the other two sides, or less than the difference of the lengths of the other two sides (Figure 1.2). The algebraic verification of these inequalities is left to the reader.

Among all points whose absolute value is the same as that of $z = (x, y)$, there is one which plays a special role. The point $(x, -y)$ is called the *conjugate* of z and is denoted by $\overline{z}$. If we view the real axis as a two-way mirror, then $\overline{z}$ is the mirror image of z (Figure 1.4).

From the definitions we obtain the following properties of the conjugate:

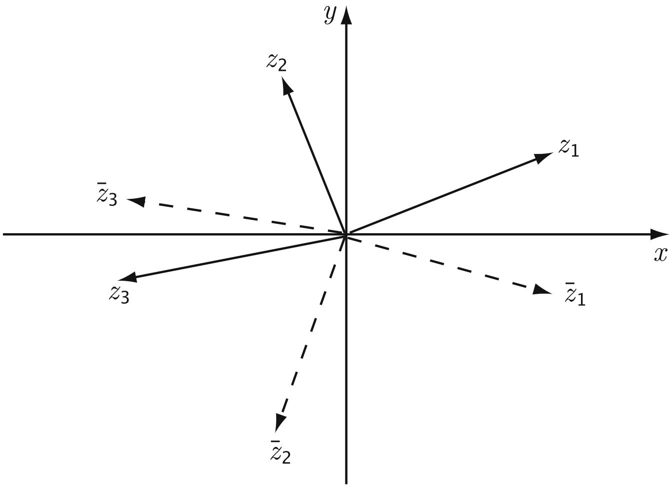


Figure 1.4. Mirror image of complex numbers

$$\begin{cases} \overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2, \\ \overline{z_1 z_2} = \overline{z}_1 \overline{z}_2. \end{cases} \tag{1.5}$$

Some of the important relationships between a complex number $z = (x, y)$ and its conjugates are

$$\begin{cases} z + \overline{z} = (2x, 0) = 2\mathrm{Re}\, z, \\ z - \overline{z} = (0, 2y) = 2i\mathrm{Im}\, z, \\ |z| = |\overline{z}| = \sqrt{x^2 + y^2}, \\ z\overline{z} = |z|^2. \end{cases} \tag{1.6}$$

The *squared* form of the absolute value in (1.6) is often the most workable. For example, to prove that the absolute value of the product of two complex numbers is equal to the product of their absolute values, we write

$$|z_1 z_2|^2 = (z_1 z_2)\overline{(z_1 z_2)} = (z_1 z_2)(\overline{z}_1 \overline{z}_2) = (z_1 \overline{z}_1)(z_2 \overline{z}_2) = (|z_1|\,|z_2|)^2.$$

Moreover, the conjugate furnishes us with a method of separating the inverse of a complex number into its real and imaginary parts:

$$(a + bi)^{-1} = \frac{1}{a + bi} \cdot \frac{\overline{a + bi}}{\overline{a + bi}} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i.$$

Equation of a line in $\mathbb{C}$. Now we may rewrite the equation of a straight line in the plane, with the real and imaginary axes as axes of coordinates, as

$$ax + by + c = 0,\ a, b, c \in \mathbb{R}; \quad \text{i.e., } a\left(\frac{z + \overline{z}}{2}\right) + b\left(\frac{z - \overline{z}}{2i}\right) + c = 0,$$

where at least one of a, b is nonzero. That is,

$$(a - ib)z + (a + ib)\overline{z} + 2c = 0.$$

Conversely, by retracing the steps above, we see that

$$\alpha z + \beta\overline{z} + \gamma = 0 \tag{1.7}$$

represents a straight line provided $\alpha = \overline{\beta}$, $\alpha \neq 0$ and γ is real.

Equation of a circle in $\mathbb{C}$. A circle in $\mathbb{C}$ is the set of all point equidistant from a given point, the center. The standard equation of a circle in the xy plane with center at (a, b) and radius $r > 0$ is $(x - a)^2 + (y - b)^2 = r^2$. If we transform this by means of the substitution $z = x + iy,\ z_0 = a + ib$, then we have $z - z_0 = (x - a) + i(y - b)$ so that

$$(z - z_0)(\overline{z - z_0}) = |z - z_0|^2 = (x - a)^2 + (y - b)^2 = r^2.$$

Therefore, the equation of the circle in the complex form with center z_0 and radius r is $|z - z_0| = r$. In complex notation we may rewrite this as

$$z\overline{z} - (z\overline{z}_0 + \overline{z}z_0) + z_0\overline{z}_0 = r^2, \quad \text{i.e.} \quad z\overline{z} - 2\mathrm{Re}\,[z(a - ib)] + a^2 + b^2 - r^2 = 0,$$

where $z_0 = a + ib$. Thus, in general, writing $a - ib = \beta$ and $\gamma = a^2 + b^2 - r^2$, we see that

$$\alpha|z|^2 + \beta z + \overline{\beta}\overline{z} + \gamma = 0, \quad \text{i.e.} \quad \left|z + \frac{\beta}{\alpha}\right|^2 = \frac{|\beta|^2 - \alpha\gamma}{\alpha^2}, \tag{1.8}$$

represents a circle provided α, γ are real, $\alpha \neq 0$ and $|\beta|^2 - \alpha\gamma > 0$.

The formulas in (1.6) produce

$$|z_1 + z_2|^2 = |z_1|^2 + 2\mathrm{Re}\,(z_1\overline{z}_2) + |z_2|^2. \tag{1.9}$$

Also, for two complex numbers z_1 and z_2, we have

(i) $|1 - \overline{z}_1 z_2|^2 - |z_1 - z_2|^2 = (1 + |z_1|\,|z_2|)^2 - (|z_1| + |z_2|)^2$, since[2]

$$\begin{aligned}
\text{L.H.S} &= (1 - \overline{z}_1 z_2)(1 - z_1\overline{z}_2) - (z_1 - z_2)(\overline{z}_1 - \overline{z}_2) \\
&= 1 - (\overline{z}_1 z_2 + z_1\overline{z}_2) + |z_1 z_2|^2 \\
&\qquad\qquad - (|z_1|^2 + |z_2|^2 - z_1\overline{z}_2 - \overline{z}_1 z_2) \\
&= 1 + |z_1 z_2|^2 - (|z_1|^2 + |z_2|^2) \\
&= (1 - |z_1|^2)(1 - |z_2|^2) \\
&= \text{R.H.S.}
\end{aligned}$$

Further, it is also clear from (i) that if $|z_1| < 1$ and $|z_2| < 1$, then

[2] L.H.S is to mean left-hand side and R.H.S is to mean right-hand side.

$$|z_1 - z_2| < |1 - z_1\overline{z}_2|$$

and if either $|z_1| = 1$ or $|z_2| = 1$, then

$$|z_1 - z_2| = |1 - \overline{z}_1 z_2|.$$

(ii) $|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2(|z_1|^2 + |z_2|^2)$ (*Parallelogram identity*); for,

$$\begin{aligned} \text{L.H.S} &= (z_1 + z_2)(\overline{z}_1 + \overline{z}_2) + (z_1 - z_2)(\overline{z}_1 - \overline{z}_2) \\ &= [|z_1|^2 + (z_1\overline{z}_2 + \overline{z}_1 z_2) + |z_2|^2] \\ &\qquad + [|z_1|^2 - (z_1\overline{z}_2 + \overline{z}_1 z_2) + |z_2|^2] \\ &= \text{R.H.S.} \end{aligned}$$

Example 1.3. Let us use the triangle inequality to find upper and lower bounds for $|z^4 - 3z + 1|^{-1}$ whenever $|z| = 2$. To do this, we need to find m and M so that $m \leq |z^4 - 3z + 1|^{-1} \leq M$ for $|z| = 2$. As $|3z - 1| \leq 3|z| + 1 = 7$ for $|z| = 2$, we have

$$|z^4 - 3z + 1| \geq \big|\,|z^4| - |3z - 1|\,\big| \geq 2^4 - 7 = 9$$

and $|z^4 - 3z + 1| \leq |z|^4 + |3z - 1| = 2^4 + 7 = 23$. Thus, for $|z| = 2$, we have

$$\frac{1}{23} \leq |z^4 - 3z + 1|^{-1} \leq \frac{1}{9}. \qquad \bullet$$

Example 1.4. Suppose that we wish to find all circles that are orthogonal to both $|z| = 1$ and $|z - 1| = 4$. To do this, we consider two circles:

$$C_1 = \{z : |z - \alpha_1| = r_1\}, \quad C_2 = \{z : |z - \alpha_2| = r_2\}.$$

These two circles are orthogonal to each other if (see Figure 1.5)

$$r_1^2 + r_2^2 = |\alpha_1 - \alpha_2|^2.$$

In view of this observation, the conditions for which a circle $|z - \alpha| = R$ is orthogonal to both $|z| = 1$ and $|z - 1| = 4$ are given by

$$1 + R^2 = |\alpha - 0|^2 \quad \text{and} \quad 4^2 + R^2 = |\alpha - 1|^2 = 1 + |\alpha|^2 - 2\text{Re}\,\alpha$$

which give $R = (|\alpha|^2 - 1)^{1/2}$ and $\text{Re}\,\alpha = -7$. Consequently,

$$\alpha = -7 + ib \quad \text{and} \quad R = (49 + b^2 - 1)^{1/2} = (48 + b^2)^{1/2}$$

and the desired circles are given by

$$C_b : |z - (-7 + ib)| = (48 + b^2)^{1/2}, \quad b \in \mathbb{R}. \qquad \bullet$$

Example 1.5. We wish to show that triangle $\triangle ABC$ with vertices z_1, z_2, z_3 is equilateral if and only if

$$z_1^2 + z_2^2 + z_3^2 = z_1 z_2 + z_2 z_3 + z_3 z_1. \tag{1.10}$$

To do this, we let $\alpha = z_2 - z_1$, $\beta = z_3 - z_2$, and $\gamma = z_1 - z_3$ so that $\alpha + \beta + \gamma = 0$. Further, if $\triangle ABC$ is equilateral, then (see Figure 1.6)

$$\begin{aligned}
\alpha + \beta + \gamma = 0 &\iff \overline{\alpha} + \overline{\beta} + \overline{\gamma} = 0 \\
&\iff \frac{\alpha\overline{\alpha}}{\alpha} + \frac{\beta\overline{\beta}}{\beta} + \frac{\gamma\overline{\gamma}}{\gamma} = 0 \\
&\iff \frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} = 0 \quad (\because |\alpha| = |\beta| = |\gamma|) \\
&\iff \frac{1}{z_2 - z_1} + \frac{1}{z_3 - z_2} + \frac{1}{z_1 - z_3} = 0 \\
&\iff (z_3 - z_2)(z_1 - z_3) + (z_2 - z_1)(z_1 - z_3) \\
&\qquad\qquad + (z_2 - z_1)(z_3 - z_2) = 0 \\
&\iff z_1^2 + z_2^2 + z_3^2 = z_1 z_2 + z_2 z_3 + z_3 z_1.
\end{aligned}$$

Conversely, suppose that (1.10) holds. Then

$$\begin{aligned}
\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} = 0 &\implies \alpha\beta + \beta\gamma + \gamma\alpha = 0 \\
&\implies \alpha\beta + \gamma(-\gamma) = 0, \quad \text{since } \alpha + \beta = -\gamma, \\
&\implies \alpha\beta = \gamma^2.
\end{aligned}$$

Thus, $\alpha\beta = \gamma^2$. Similarly, $\beta\gamma = \alpha^2$ and $\gamma\alpha = \beta^2$. Further,

$$(\alpha\beta)(\overline{\alpha\beta}) = \gamma^2(\overline{\gamma})^2, \quad \text{i.e.,} \quad (\alpha\overline{\alpha})(\beta\overline{\beta})(\gamma\overline{\gamma}) = (\gamma\overline{\gamma})^3.$$

Because of the symmetry, we also have

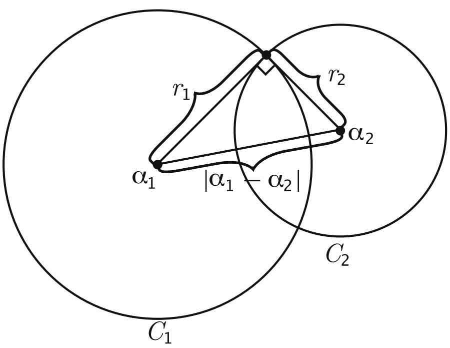


Figure 1.5. Orthogonal circles

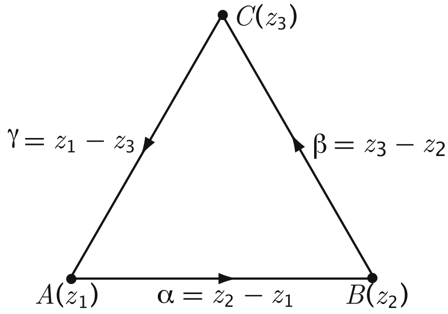


Figure 1.6. Equilateral triangle $\triangle ABC$

$$(\alpha\overline{\alpha})(\beta\overline{\beta})(\gamma\overline{\gamma}) = (\alpha\overline{\alpha})^3 \quad \text{and} \quad (\alpha\overline{\alpha})(\beta\overline{\beta})(\gamma\overline{\gamma}) = (\beta\overline{\beta})^3.$$

Thus,

$$\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} = 0 \Longrightarrow |\alpha|^3 = |\beta|^3 = |\gamma|^3 \Longrightarrow |\alpha| = |\beta| = |\gamma|,$$

showing that $\triangle ABC$ is equilateral.

Here is an alternate proof. First we remark that equilateral triangles are preserved under linear transformations $f(z) = az + b$, which can be easily verified by replacing z_j by $az_j + b$ $(j = 1, 2, 3)$ in (1.10). By a suitable transformation, we can reduce the problem to a simpler one. If z_1, z_2, z_3 are the vertices of a degenerated equilateral triangle (i.e., $z_1 = z_2 = z_3$), then (1.10) holds. If two of the vertices are distinct, then, by a suitable transformation, we can take $z_1 = 0$ and $z_2 = 1$. Then (1.10) takes the form $1 + z_3^2 = z_3$, which gives

$$z_3 = \frac{1 + i\sqrt{3}}{2} \quad \text{or} \quad \frac{1 - i\sqrt{3}}{2}.$$

In either case $\{0, 1, z_3\}$ forms vertices of an equilateral triangle. ●

Example 1.6. Suppose we wish to describe geometrically the set S given by

$$S = \{z : |z - a| - |z + a| = 2c\} \quad (0 \neq a \in \mathbb{C},\ c \geq 0), \tag{1.11}$$

for the following situations:

$$\text{(i) } c > |a| \qquad \text{(ii) } c = 0 \qquad \text{(iii) } 0 < c < a \qquad \text{(iv) } c = a > 0.$$

The triangle inequality gives that

$$|2a| = |z - a - (z + a)| \geq |z - a| - |z + a| = 2c, \quad \text{i.e., } c \leq |a|.$$

Thus, there are no complex numbers satisfying (1.11) if $c > |a|$. Hence, $S = \emptyset$ whenever $c > |a|$.

If $c = 0$, we have $|z - a| = |z + a|$ which shows that S is the line that is the perpendicular bisector of the line joining a and $-a$.

Next, we consider the case $a > c > 0$. Then, writing $z = x + iy$,

$$\begin{aligned} |z-a| - |z+a| = 2c &\iff |z-a|^2 = (2c + |z+a|)^2 \\ &\iff |z-a|^2 = 4c^2 + |z+a|^2 + 4c|z+a| \\ &\iff c|z+a| + c^2 = -a\mathrm{Re}\, z \quad (\mathrm{Re}\, z < 0) \\ &\iff c^2[|z|^2 + a^2 + 2a\mathrm{Re}\, z] = (c^2 + a\mathrm{Re}\, z)^2 \\ &\iff c^2|z|^2 - a^2(\mathrm{Re}\, z)^2 = c^2(c^2 - a^2) \\ &\iff \frac{x^2}{c^2} - \frac{y^2}{a^2 - c^2} = 1. \end{aligned}$$

Further, we observe that for $|z - a| - |z + a|$ to be positive, we must have $\mathrm{Re}\, z < 0$. Thus, if $a > c > 0$ we have

$$S = \left\{ x + iy : \frac{x^2}{c^2} - \frac{y^2}{a^2 - c^2} = 1 \right\}$$

and so S describes a hyperbola with focii at $a, -a$.

Finally, if $c = a$ then

$$|z-a| - |z+a| = 2a \iff |z+a| = -\mathrm{Re}\,(z+a) \implies \mathrm{Re}\,(z+a) < 0$$

and therefore, S in this case is the interval $(-\infty, -a]$. ●

Questions 1.7.

1. In Figure 1.2, would we still have a parallelogram if the vector z_2 were in the same or the opposite direction as that of z_1?
2. Geometrically, can we predict the quadrant of $z_1 + z_2$ from our knowledge of z_1 and z_2?
3. Why don't we define multiplication of complex numbers as vector multiplication?
4. When does the triangle inequality become an equality?
5. What would be the geometric interpretation of the inequality for the sum of n complex numbers?
6. Name some interesting relationships between the points (x, y) and $(-x, y)$.
7. If a and b are positive rational numbers, why might we want to call the numbers $\sqrt{a} + \sqrt{b}$ and $\sqrt{a} - \sqrt{b}$ real conjugates?
8. Is every rational number algebraic? Are $\sqrt{3}$ and $\sqrt[5]{5} - 3i$ algebraic?
   Note: A number is algebraic if it is a solution of a polynomial (in z) with integer coefficients. Numbers which are not algebraic are called transcendental numbers.
9. What does $|z|^2 + \beta z + \overline{\beta}\overline{z} + \gamma = 0$ represent if $|\beta|^2 \geq \gamma$?
10. Is $|z + 1| + |z - 1| \leq 2\sqrt{2}$ if $|z| \leq 1$?

Exercises 1.8.

1. If $z_1 = 3 - 4i$ and $z_2 = -2 + 3i$, obtain graphically and analytically
   (a) $2z_1 + 4z_2$ (b) $3z_1 - 2\overline{z}_2$ (c) $z_1 - \overline{z}_2 - 4$
   (d) $|z_1 + z_2|$ (e) $|z_1 - z_2|$ (f) $|2\overline{z}_1 + 3\overline{z}_2 - 1|$.
2. Let $z_1 = x_1 + iay_1$ and $z_2 = x_2 - ib/y_1$, where a, b are real. Determine a condition on y_1 so that $z_1^{-1} + z_2^{-1}$ is real.
3. Identify all the points in the complex plane that satisfy the following relations.
   (a) $1 < |z| \leq 3$ (b) $|(z-3)/(z+3)| < 2$
   (c) $|z-1| + |z+1| = 2$ (d) $\mathrm{Re}\,(z-5) = |z| + 5$
   (e) $\mathrm{Re}\, z^2 > 0$ (f) $\mathrm{Im}\, z^2 > 0$
   (g) $\mathrm{Re}\,((1-i)z) = 2$ (h) $|z - i| = \mathrm{Re}\, z$
   (i) $\mathrm{Re}\,(z) = |z|$ (j) $\mathrm{Re}\,(z^2) = 1$
   (k) $\overline{z} = 5/(z-1)$ $(z \neq 1)$ (l) $[\mathrm{Im}\,(iz)]^2 = 1$.
4. Let $|(z-a)/(z-b)| = M$, where a and b are complex constants and $M > 0$. Describe this curve and explain what happens as $M \to 0$ and as $M \to \infty$.
5. Find a complex form for the hyperbola with real equation $9x^2 - 4y^2 = 36$.
6. If $|z| < 1$, prove that

$$\text{(a) Re}\left(\frac{1}{1-z}\right) > \frac{1}{2} \quad \text{(b) Re}\left(\frac{z}{1-z}\right) > -\frac{1}{2} \quad \text{(c) Re}\left(\frac{1+z}{1-z}\right) > 0.$$

7. If $P(z)$ is a polynomial equation with real coefficients, show that z_1 is a root if and only if $\overline{z}_1$ is a root. Conclude that any polynomial equation of odd degree with real coefficients must have at least one real root. Can you prove this using elementary calculus?
8. Prove that, for every $n \geq 1$,

$$|z_1 + z_2 + \cdots + z_n| \leq |z_1| + |z_2| + \cdots + |z_n|.$$

9. Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be complex numbers. Prove the Schwarz inequality,

$$\left|\sum_{k=1}^{n} a_k b_k\right|^2 \leq \left(\sum_{k=1}^{n} |a_k|^2\right)\left(\sum_{k=1}^{n} |b_k|^2|\right).$$

   When will equality hold?
10. Define $e(\alpha) = \cos\alpha + i\sin\alpha$, for α real. Prove the following.
    (a) $e(0) = 1$ (b) $|e(\alpha)| = 1$
    (c) $e(\alpha_1 + \alpha_2) = e(\alpha_1)e(\alpha_2)$ (d) $e(n\alpha) = [e(\alpha)]^n$.

    Which of these properties does the real-valued function $f(x) = e^x$ satisfy?

11. Show that the line connecting the complex numbers z_1 and z_2 is perpendicular to the line connecting z_3 and z_4 if and only if

$$\mathrm{Re}\,\{(z_1 - z_2)(\overline{z}_3 - \overline{z}_4)\} = 0.$$

12. If a, b are real numbers in the unit interval $(0, 1)$, then when do the three points $z_1 = a + i$, $z_2 = 1 + ib$ and $z_3 = 0$ form an equilateral triangle?
13. If $|z_j| = 1$ $(j = 1, 2, 3)$ such that $z_1 + z_2 + z_3 = 0$, then show that z_j's are the vertices of an equilateral triangle.

1.3 Polar Representation

In Section 1.2, the magnitude of the vector $z = x + iy$ was discussed. What about its direction? A measurement of the angle θ that the vector $z\,(\neq 0)$ makes with the positive real axis is called an *argument* of z (see Figure 1.7). Thus, we may express the point $z = (x, y)$ in the "new" form

$$(r\cos\theta, r\sin\theta).$$

This, of course, is just the polar coordinate representation for the complex number z. We have the familiar relations

$$r = |z| = \sqrt{x^2 + y^2} \quad \text{and} \quad \tan\theta = \frac{y}{x}.$$

The real numbers r and θ, like x and y, uniquely determine the complex number z. Unfortunately, the converse isn't completely true. While z uniquely determines the x and y, hence r, the value of θ is determined up to a multiple of 2π. There are infinitely many distinct arguments for a given complex number z, and the symbol $\arg z$ is used to indicate *any one* of them. Thus the arguments of the complex number (2, 2) are

$$\frac{\pi}{4} + 2k\pi \quad (k = 0, \pm 1, \pm 2, \dots).$$

This inconvenience can sometimes (although not always) be ignored by distinguishing (arbitrarily) one particular value of $\arg z$. We use the symbol $\mathrm{Arg}\, z$ to stand for the unique determination of θ for which $-\pi < \arg z \leq \pi$. This θ is called the *principal value* of the argument. To illustrate,

$$\mathrm{Arg}\,(2,2) = \frac{\pi}{4}, \quad \mathrm{Arg}\,(0,-5) = -\frac{\pi}{2}, \quad \mathrm{Arg}\,(-1,\sqrt{3}) = \frac{2\pi}{3}.$$

Note that $\mathrm{Re}\, z > 0$ is equivalent to $|\mathrm{Arg}\, z| < \pi/2$. If $x = y = 0$, the expression $\tan\theta = y/x$ has no meaning. For this reason, $\arg z$ is not defined when $z = 0$.

Suppose that z_1 and z_2 have the polar representations

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

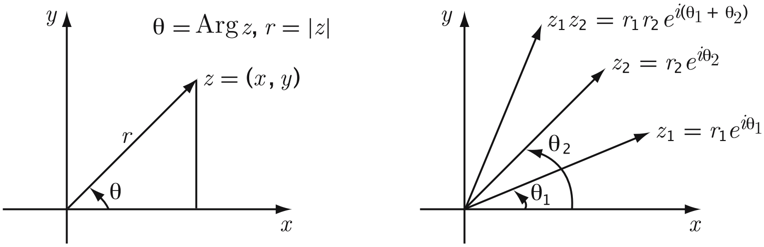


Figure 1.7. Polar representation of z and $z_1 z_2$

Then

$$\begin{aligned} z_1 z_2 &= r_1 r_2[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \sin\theta_2\cos\theta_1)] \\ &= r_1 r_2[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)]. \end{aligned}$$

Loosely speaking, we may say that the argument of the product of two nonzero complex numbers is equal to the sum of their arguments; that is,

$$\arg(z_1 z_2) = \arg z_1 + \arg z_2. \tag{1.12}$$

We understand (1.12) to mean that if θ_1 is one of the values of $\arg z_1$ and θ_2 is one of the values of $\arg z_2$, then $\theta_1 + \theta_2$ is one of the values of $\arg(z_1 z_2)$. Since (1.12) is valid only up to a multiple of 2π, a more explicit formulation is

$$\arg z_1 z_2 = \arg z_1 + \arg z_2 + 2k\pi \quad (k \text{ an integer})$$

or

$$\arg z_1 z_2 = \arg z_1 + \arg z_2 \pmod{2\pi}$$

(see Figure 1.7). To illustrate, we observe that if $z = (-1 + i\sqrt{3})/2$, then $z^2 = (-1 - i\sqrt{3})/2$ so that

$$\operatorname{Arg} z = \frac{2\pi}{3} \ \text{ and } \ \operatorname{Arg}(z^2) = -\frac{2\pi}{3}.$$

Thus, $\operatorname{Arg}(z.z) = \operatorname{Arg} z + \operatorname{Arg} z - 2\pi$.

An induction argument (no pun intended) shows that if z_i has modulus r_i and argument θ_i $(i = 1, 2, \dots, n)$, then

$$\begin{aligned} z_1 z_2 \cdots z_n = r_1 r_2 \cdots r_n[&\cos(\theta_1+\theta_2+\cdots+\theta_n) \\ &+ i\sin(\theta_1+\theta_2+\cdots+\theta_n)]. \end{aligned} \tag{1.13}$$

Example 1.9. Let $z_1 = 1 + i$ and $z_2 = \sqrt{3} + i$. We wish to express them in polar form and then verify the identities that hold for multiplication and division of z_1 and z_2, respectively. To do this, we may write

$$z_1 = \sqrt{2}e^{i\pi/4} \ \text{ and } \ z_2 = 2e^{i\pi/6}.$$

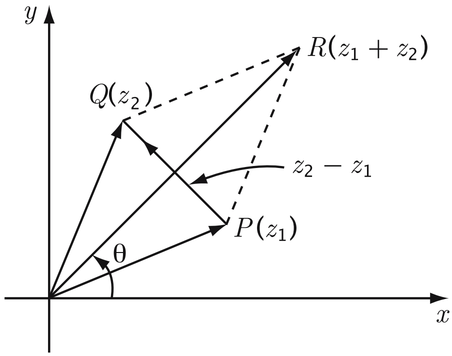


Figure 1.8. Geometric proof for Example 1.10

Then

$$z_1 z_2 = 2\sqrt{2}e^{i5\pi/12} \text{ and } \frac{z_1}{z_2} = \frac{1}{\sqrt{2}}e^{i\pi/12}.$$

Thus, in this particular problem of product and division, it follows that

$$\mathrm{Arg}\,(z_1 z_2) = \mathrm{Arg}\, z_1 + \mathrm{Arg}\, z_2 \text{ and } \mathrm{Arg}\left(\frac{z_1}{z_2}\right) = \mathrm{Arg}\, z_1 - \mathrm{Arg}\, z_2.$$

Similarly, we may easily check the following:

(i) $(1 - i\sqrt{3})/(1 + i\sqrt{3}) = e^{i\theta},\ \theta = 2\pi/3 + 2k\pi$;
(ii) $(-\sqrt{3} + i)(1 + i)/(1 + i\sqrt{3}) = \sqrt{2}e^{i\theta},\ \theta = 3\pi/4 + 2k\pi$;
(iii) $(1 - 3i)/(2 - i) = \sqrt{2}e^{i\theta},\ \theta = -\pi/4 + 2k\pi$,

where k is an integer. ●

Example 1.10. Suppose that z_1 and z_2 are two nonzero complex numbers such that $|z_1| = |z_2|$ but $z_1 \neq \pm z_2$. Then we wish to show that the quotient $(z_1 + z_2)/(z_1 - z_2)$ is a purely imaginary number. For a geometric proof, we consider the parallelogram $OPRQ$ shown in Figure 1.8. Since the sides OP and OQ are equal in length, $OPRQ$ is a rhombus. Thus, the vector $\overrightarrow{OR}$ is perpendicular to the vector $\overrightarrow{PQ}$, and so

$$\mathrm{Arg}\,(z_1 + z_2) = \mathrm{Arg}\,(z_1 - z_2) \pm i\pi/2.$$

For an analytic proof, we may rewrite

$$w = \frac{z_1 + z_2}{z_1 - z_2} = \frac{1 + z}{1 - z} \quad (z = z_2/z_1).$$

The hypotheses imply that $|z| = 1,\ z \neq \pm 1$. Therefore, letting $z = e^{i\theta}$ with $\theta \in (0, 2\pi)\backslash\{\pi\}$,

$$w = \frac{1 + e^{i\theta}}{1 - e^{i\theta}} = \frac{e^{-i\theta/2} + e^{i\theta/2}}{e^{-i\theta/2} - e^{i\theta/2}} = \frac{2\cos(\theta/2)}{-2i\sin(\theta/2)} = i\cot(\theta/2),$$

which is a purely imaginary number. ●

Example 1.11. Let $z = \sin\theta + i\cos 2\theta$ and $w = \cos\theta + i\sin 2\theta$. We wish to show that there exists no value of θ for which $z = w$. To do this, we first note that

$$z = w \Longleftrightarrow \sin\theta = \cos\theta \quad \text{and} \quad \cos 2\theta = \sin 2\theta.$$

There exists no values of θ satisfying both conditions, because $\sin\theta = \cos\theta$ implies that $\cos 2\theta = \cos^2\theta - \sin^2\theta = 0$, and so the second condition reduces to $\sin 2\theta = 2\sin\theta\cos\theta = 0$, i.e., $\sin\theta = 0 = \cos\theta$. ●

Remark 1.12. Geometric considerations (Figures 1.2 and 1.7) indicate that the rectangular representation will frequently be more useful for problems involving sums of complex numbers, with polar representation being more useful for problems involving products. ●

If we let $z_1 = z_2 = \cdots = z_n$ in (1.13), we obtain

$$z^n = r^n(\cos n\theta + i\sin n\theta). \tag{1.14}$$

For $|z| = 1$ (the unit circle), (1.14) reduces to

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta, \tag{1.15}$$

a theorem of DeMoivre.

The possibility of finding nth roots of the complex number is suggested by (1.14). A complex number z is an nth root of z_0 if $z^n = z_0$, written $z = z_0^{1/n}$.

The problem is to reverse the multiplicative operation and determine a number which, when multiplied by itself n times, furnishes us with the original number. Given a complex number $z_0 = r_0(\cos\theta_0 + i\sin\theta_0)$, how do you find a complex number $z = r(\cos\theta + i\sin\theta)$ such that $z^n = z_0$? By (1.14), we must have

$$r^n(\cos n\theta + i\sin n\theta) = r_0(\cos\theta_0 + i\sin\theta_0). \tag{1.16}$$

Since $|\cos\alpha + i\sin\alpha| = 1$ for all real α, (1.16) yields the relations

$$r^n = r_0, \quad \cos n\theta + i\sin n\theta = \cos\theta_0 + i\sin\theta_0. \tag{1.17}$$

The first relation in (1.17) shows that $|z| = r_0^{1/n}$, which we already knew (why)? But the second gives important information about the argument of z, namely, that $n\arg z$ differs from $\arg z_0$ by a multiple of 2π (that is, $n\theta = \theta_0 + 2k\pi$, $k = 0, \pm 1, \pm 2, \dots$):

$$\theta = \frac{\theta_0 + 2k\pi}{n}. \tag{1.18}$$

How many integers k in (1.18) produce distinct solutions? We have

$$z = z_0^{1/n} = r_0^{1/n}\left\{\cos\left(\frac{\theta_0 + 2k\pi}{n}\right) + i\sin\left(\frac{\theta_0 + 2k\pi}{n}\right)\right\}. \tag{1.19}$$

For each k $(k = 0, 1, 2, \ldots, n-1)$, there is a different value for z. We leave it for the reader to verify that there are no more solutions. Thus, given $z_0 \neq 0$, there are exactly n distinct complex numbers z such that $z^n = z_0$.

By letting $z_0 = 1$ in (1.19), we may find the nth roots of unity. If $z^n = 1$, then

$$z = \cos\left(\frac{2k\pi}{n}\right) + i\sin\left(\frac{2k\pi}{n}\right) \quad (k = 0, 1, 2, \ldots, n-1). \tag{1.20}$$

Geometrically, the solutions represent the n vertices of a regular polygon of n sides inscribed in a circle with center at the origin and radius equal to one. See Figures 1.9 and 1.10 for the inscribed square and pentagon.

By (1.20), the difference in the arguments of any two successive nth roots of unity is constant $(2\pi/n)$. If we let

$$\omega = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n},$$

then each root of unity may be expressed as a multiple of ω; that is,

$$\omega, \omega^2, \omega^3, \ldots, \omega^{n-1}, \quad \omega^n = \omega^0 = 1.$$

This gives interesting information about the sums and products of the roots of the unity, namely, that the product of any two roots of unity is also a root of unity, and that the sum of all nth roots of unity is zero. The latter statement follows from the identify

$$1 + \omega + \omega^2 + \cdots + \omega^{n-1} = \frac{1-\omega^n}{1-\omega}.$$

Using (1.19), we easily see, for instance, the following:

(a) $*\sqrt{3+4i} = \pm(2+i)$
(b) $*\sqrt{-3+4i} = \pm(1+2i)$

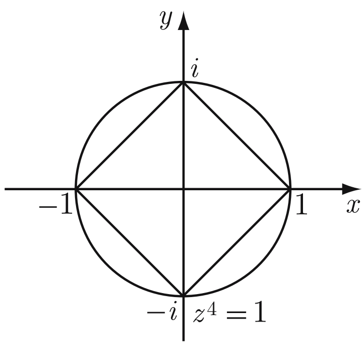


Figure 1.9. Illustration for the 4th roots of unity

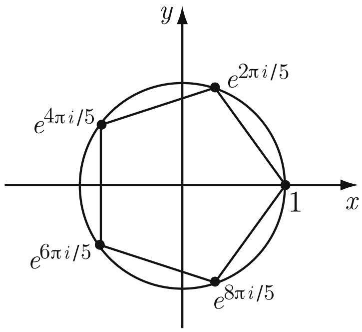


Figure 1.10. Illustration for the 5th roots of unity

(c) $*\sqrt{1+i} = \pm\left(\sqrt{\frac{\sqrt{2}+1}{2}} + i\sqrt{\frac{\sqrt{2}-1}{2}}\right)$
(d) $*\sqrt{2i} = \pm(1+i)$
(e) $*\sqrt{\frac{1-i\sqrt{3}}{2}} = \pm\left(\frac{\sqrt{3}-i}{2}\right)$
(f) $*\sqrt{1+i\sqrt{3}} = \pm\left(\frac{\sqrt{3}+i}{\sqrt{2}}\right)$
(g) $*\sqrt{-5-12i} = \pm(-2+3i)$
(h) $*\sqrt{5+12i} = \pm(3+2i)$
(i) $*\sqrt{-5+12i} = \pm(2+3i)$.

Here $*\sqrt{a+ib}$ denotes the two 2th roots of the complex number $a+ib$.

Since the n nth roots of unity are given by (1.20), we have

$$z^n - 1 = (z-1)(z-\omega_1)(z-\omega_2)\cdots(z-\omega_{n-1}), \quad \omega_k = \omega^k = e^{2\pi ki/n}.$$

Dividing both sides by $z-1$, using the identity

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

and letting $z \to 1$, we have

$$n = (1-\omega_1)(1-\omega_2)\cdots(1-\omega_{n-1}), \text{ and}$$
$$n = (1-\overline{\omega}_1)(1-\overline{\omega}_2)\cdots(1-\overline{\omega}_{n-1}).$$

As $(1-e^{-i\theta})(1-e^{i\theta}) = 2(1-\cos\theta) = 4\sin^2(\theta/2)$, it follows that

$$n^2 = \prod_{k=1}^{n-1}|1-\omega_k|^2 = \prod_{k=1}^{n-1}\left\{4\sin^2\left(\frac{k\pi}{n}\right)\right\} = 2^{2(n-1)}\prod_{k=1}^{n-1}\sin^2\left(\frac{k\pi}{n}\right).$$

Taking the positive square root on both sides we have

$$n = 2^{n-1} \prod_{k=1}^{n-1} \sin\left(\frac{k\pi}{n}\right), \quad n > 1. \tag{1.21}$$

We can make the following generalization: Consider the equation

$$M_a(z) = z^{2n} - 2z^n a^n \cos n\phi + a^{2n} = 0 \quad (n \in \mathbb{N},\ a \in \mathbb{R}^+,\ \phi \in \mathbb{R}).$$

Solving this for z^n, we find $z^n = a^n e^{\pm in\phi}$ so that

$$M_a(z) = [z^n - a^n e^{in\phi}][z^n - a^n e^{-in\phi}].$$

Therefore, using the concept of nth root of a complex number, we can write

$$\begin{aligned} M_a(z) &= \prod_{k=1}^{n} \left[z - ae^{i(\phi+2k\pi/n)}\right]\left[z - ae^{-i(\phi+2k\pi/n)}\right] \\ &= \prod_{k=1}^{n} \left[z^2 - 2za\cos\left(\phi + \frac{2k\pi}{n}\right) + a^2\right]. \end{aligned} \tag{1.22}$$

Some special cases of (1.22) follow:

(a) Taking $\phi = 0$, we have

$$(z^n - a^n)^2 = \prod_{k=1}^{n} \left[z^2 - 2za\cos\left(\frac{2k\pi}{n}\right) + a^2\right].$$

(b) Taking $\phi = \pi/n$, we have

$$(z^n + a^n)^2 = \prod_{k=1}^{n} \left[z^2 - 2za\cos\left(\frac{(2k+1)\pi}{n}\right) + a^2\right].$$

(c) If $a = 1$ then, on dividing (1.22) by z^n, $z \neq 0$, we have

$$z^n + z^{-n} - 2\cos(n\phi) = \prod_{k=1}^{n} \left[z + z^{-1} - 2\cos\left(\phi + \frac{2k\pi}{n}\right)\right]$$

and so, if $z = e^{i\theta}$, this becomes

$$\cos(n\theta) - \cos(n\phi) = 2^{n-1} \prod_{k=1}^{n} \left[\cos\theta - \cos\left(\phi + \frac{2k\pi}{n}\right)\right]$$

which is, for $\cos\theta \neq \cos\phi$, equivalent to

$$\frac{\cos(n\theta) - \cos(n\phi)}{\cos\theta - \cos\phi} = 2^{n-1} \prod_{k=1}^{n-1} \left[\cos\theta - \cos\left(\phi + \frac{2k\pi}{n}\right)\right].$$

In the limiting case when $\theta, \phi \to 0$, the above reduces to

$$n = 2^{n-1} \prod_{k=1}^{n-1} \sin\left(\frac{k\pi}{n}\right),$$

which is nothing but (1.21).

Questions 1.13.

1. What problem would be created by defining the argument of $z = 0$ to be zero?
2. Loosely speaking, for complex numbers z_1 and z_2 we have

$$\arg(z_1 z_2) = \arg z_1 + \arg z_2.$$

What real-valued functions have the property that

$$f(x_1 x_2) = f(x_1) + f(x_2)?$$

3. When does $\operatorname{Arg}(z_1 z_2) = \operatorname{Arg} z_1 + \operatorname{Arg} z_2$?
4. How are the complex numbers z_1 and z_2 related if $\arg(z_1) = \arg z_2$?
5. How are the arguments $\arg(z_1)$ and $\arg z_2$ related if $z_1 = z_2$?
6. How are the arguments $\arg(z_1)$ and $\arg z_2$ related if $\operatorname{Re}(z_1 \overline{z}_2) = |z_1 z_2|$?
7. How are the arguments $\arg(z_1)$ and $\arg z_2$ related if $|z_1 + z_2| = |z_1| + |z_2|$?
8. As the complex number z approaches the negative real axis from above and below, what is happening to $\operatorname{Arg} z$? What if z approaches the *positive* real axis from above and below?
9. How do the arguments of z and $1/z$ compare?
10. How do the arguments of z and $\overline{z}$ compare?
11. How do the arguments of $\overline{z}$ and $1/z$ compare?
12. What is the position of the complex number $(\cos\alpha + i \sin\alpha)z$ relative to the position of z?
13. What are some differences between the terms angle, real number, and argument?
14. Of what use might the binomial theorem be in this section?
15. For which integers n does $z^n = 1$ have only real solutions?
16. For which complex numbers z does $\sqrt{z/\overline{z}} = z/|z|$?
17. Is it always the case that for any given nonzero complex number, either $\sqrt{z^2} = z$ or $\sqrt{z^2} = -z$?
18. Which postulates for a field are satisfied by the roots of unity under ordinary addition and multiplication of complex numbers?
19. What can you say about the nth roots of an arbitrary complex number?
20. For α an arbitrary real number, how many solutions might you expect $z^\alpha = 1$ to have?
21. If $z = e^{i\alpha}$ ($\alpha \in (0, 2\pi)$), is $(1 + z)/(1 - z)$ equal to $i \cot(\alpha/2)$?

Exercises 1.14.

1. For a fixed positive integer n, determine the real part of $(1+i\sqrt{3})^n$.
2. Find two complex numbers z_1 and z_2 so that
$$\operatorname{Arg}(z_1z_2) \neq \operatorname{Arg} z_1 + \operatorname{Arg} z_2.$$
3. Find two complex numbers z_1 and z_2 so that
$$\operatorname{Arg}(z_1z_2) = \operatorname{Arg} z_1 + \operatorname{Arg} z_2.$$
4. Describe the following regions geometrically.
   (a) $\operatorname{Arg} z = \pi/6, \quad |z| > 1$ (b) $\pi/4 < \operatorname{Arg} z < \pi/2$
   (c) $-\pi < \operatorname{Arg} z < 0, \quad |z+i| > 2$ (d) $1 < |z-1| < 5$.
5. If $|1-z| < 1$, show that $|\operatorname{Arg} z| < \pi/2$.
6. If $|z| < 1$, show that $|\operatorname{Arg}((1+z)/(1-z))| < \pi/2$.
7. If $\operatorname{Re} z > 0$, show that $\operatorname{Re}(1/z) > 0$. If $\operatorname{Re} z > a > 0$, what can you say about $\operatorname{Re}(1/z)$?
8. If $|z| = 1$, $z \neq -1$, show that z may be expressed in the form
$$z = \frac{1+it}{1-it},$$
where t is a real number.
9. Write the polar form of the following:
   (a) $\dfrac{1+\cos\phi + i\sin\phi}{1+\cos\phi - i\sin\phi} \quad (0 < \phi < \pi/2)$
   (b) $\dfrac{1+\cos\phi + i\sin\phi}{1-\cos\phi - i\sin\phi}$
   (c) $1-\sin\phi + i\cos\phi \quad (0 < \phi < \pi/2)$
   (d) $-\sin\phi - i\cos\phi$
   (e) $(1+i)^n \quad (n \in \mathbb{N})$
   (f) $(1+i\sqrt{3})^n + (1-i\sqrt{3})^n \quad (n \in \mathbb{N})$.
10. Find all values of the following and simplify the expressions as much as possible.
   (a) $i^{1/2}$ (b) $i^{1/4}$ (c) $(-i)^{1/3}$ (d) $\sqrt{1+i}$
   (e) $\sqrt[6]{8}$ (f) $\sqrt{4+3i}$ (g) $(4-3i)^{1/3}$ (h) $\sqrt{2+i}$
11. If $\omega = (-1+i\sqrt{3})/2$ is a cube root of unity and if
$$S_n = 1 - \omega + \omega^2 + \cdots + (-1)^{n-1}\omega^{n-1},$$
then find a formula for S_n.
12. Let ω be a cube root of unity and let a, b, c be real. Determine a condition on a, b, c so that $(a+b\omega+c\omega^2)^3$ is real.
13. Let ω be a cube root of unity. Determine the value of
   (a) $(1+\omega)^3$ (b) $(1+2\omega+\omega^2)(1+\omega+2\omega^2)$
   (c) $(1+\omega+2\omega^2)^9$ (d) $(1+3\omega+2\omega^2)(1+4\omega+3\omega^2)$.

14. Let $\omega \neq 1$ be an nth root of unity. Show that
$$1 + 2\omega + 3\omega^2 + \cdots + n\omega^{n-1} = -\frac{n}{1-\omega}.$$
15. Let $\omega_k = \cos(2k\pi/n) + i\sin(2k\pi/n)$. Show that $\sum_{k=1}^{n} |\omega_k - \omega_{k-1}| < 2\pi$ for all values of n. What happens as n approaches ∞?
16. Find the roots of the equation $(1+z)^5 = (1-z)^5$.
17. Find α, β, γ and δ such that the roots of the equation
$$z^5 + \alpha z^4 + \beta z^3 + \gamma z^2 + \delta z + \eta = 0$$
lie on a regular pentagon centered at 1.
18. Prove that for any real x and a natural number n,
$$e^{i2n\cot^{-1}(x)} \left(\frac{ix+1}{ix-1}\right)^n = 1.$$
19. Find a positive integer n such that

    (i) $(\sqrt{3}+i)^n = 2^n$ (ii) $(-1+i)^n = 2^{n/2}$.

2

Topological and Analytic Preliminaries

The neighborhood of a young child consists of the people very close on the left and right. As we get older we think in terms of two-dimensional neighborhoods (the people around the corner) or even three-dimensional neighborhoods (the people in the world). In this chapter we do likewise. We develop numerous methods for accurately describing sets in the real line (one-dimensional) and the plane (two-dimensional). In order to track down the elusive point at infinity, it becomes necessary to introduce the sphere (three-dimensional).

When a set is described in a satisfactory manner, we become concerned about its image. We investigate conditions under which properties of a set are preserved when the set is transformed into a new set. A remarkable outcome of our investigation is that the removal of a single point from one set may entirely change its character, whereas the removal of infinitely many points from a different set may be insignificant. The removal of two points from a set on the line may give it more affinity to a set in the plane than to its former self. In this chapter we learn that in a sense all points are equal but some points are more equal than others.

2.1 Point Sets in the Plane

A neighborhood of a real number x_0 is an interval in the form $(x_0 - \epsilon, x_0 + \epsilon)$, where ϵ is any positive real number. Thus we may say that an ϵ neighborhood of x_0 is the set of points $x \in \mathbb{R}$ for which $|x - x_0| < \epsilon$. There are different ways to extend this one-dimensional neighborhood concept to include points in the plane. A *square ϵ neighborhood* of a point (x_0, y_0) is the set of all points (x, y) whose coordinates satisfy the two inequalities

$$|x - x_0| < \epsilon, \quad |y - y_0| < \epsilon.$$

It consists of all points inside a square centered at (x_0, y_0). The sides of the square are parallel to the coordinate axes and have length 2ϵ. A *circular ϵ*

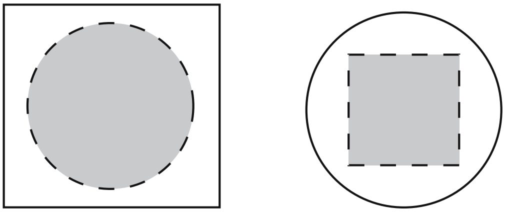


Figure 2.1. Illustration for open sets in the plane

neighborhood of (x_0, y_0) is the set of all points (x, y) whose distance from (x_0, y_0) is less than ϵ. It consists of all points (x, y) such that

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} < \epsilon,$$

i.e., points inside a circle centered at (x_0, y_0) whose radius is ϵ. Observe that every square neighborhood of a point contains a circular neighborhood of the point, and every circular neighborhood of a point contains a square neighborhood of the point (for a smaller ϵ, of course). This is illustrated in Figure 2.1. From our point of view (that a point in the plane represents a complex number), it will be more convenient to deal with circular neighborhoods, for then an ϵ neighborhood of the complex number z_0 consists of all points $z \in \mathbb{C}$ satisfying the inequality $|z - z_0| < \epsilon$. Such a neighborhood is denoted by $N(z_0; \epsilon)$.

Care must be taken to distinguish between a neighborhood on the real line and a neighborhood in the plane. For example, $\{x \in \mathbb{R} : -1 < x < 1\}$ is a neighborhood of 0, a point on the line; it is *not* a neighborhood of $(0, 0)$, a point in the plane. A point in the plane is not permitted to have a one-dimensional neighborhood.

The definitions and theorems in this section are valid simultaneously for points on the line and points in the plane, when the concepts of ϵ neighborhood are suitably interpreted. A set is said to be *bounded* if it is contained in some disk centered at the origin. A point is said to be an *interior point* of a set if there is some neighborhood of the point contained in the set. An important distinction between the bounded sets

$$A = \{z \in \mathbb{C} : |z - z_0| < \epsilon\} \quad \text{and} \quad B = \{z \in \mathbb{C} : |z - z_0| \leq \epsilon\}$$

is that every point in A is an *interior point*. To see this, let z_1 be any point in A. Then $|z - z_0| = \delta$ for some δ, $0 \leq \delta < \epsilon$. But for $\eta = (\epsilon - \delta)/2$, we have $N(z_1; \eta) \subset N(z_0; \delta)$ (see Figure 2.2). Of course, no point on the circle $|z - z_0| = \epsilon$ is an interior point of B. A set A is called an *open set* if every point in A is an interior point. We have shown that a neighborhood of a point in the plane is an open set. Other simple examples of open sets in the plane are

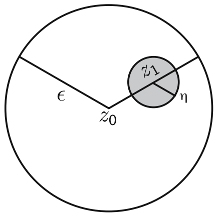


Figure 2.2. Description for an interior point

(a) the empty set,
(b) the set of all complex numbers,
(c) $\{z : |z| > r\}$, $r \geq 0$,
(d) $\{z : r_1 < |z| < r_2\}$, $0 \leq r_1 < r_2$,
(e) the intersection of any two open sets,
(f) the union of any collection of open sets.

Remark 2.1. An open interval on the real line is not an open set in the plane, since any neighborhood of a point will contain points in the plane that are not real. ●

A *deleted* ϵ neighborhood of z_0, denoted by $N'(z_0; \epsilon)$, is the set of all points z such that $0 < |z - z_0| < \epsilon$. That is, the point z_0 is "punched out" from the set. A point z_0 is called a *limit point* of a set A if every deleted neighborhood of z_0 contains a point of A. Note that a limit point z_0 may or may not be in the set A.

Examples 2.2. (i) The limit points of the open set $|z| < 1$ are $|z| \leq 1$; that is, all the points of the set and all the points on the unit circle $|z| = 1$. If $\partial \Delta = \{z : |z| = 1\}$ and $\overline{\Delta} = \{z : |z| \leq 1\}$, then all points of $\overline{\Delta}$ are its limit points and no other point is a limit point of $\overline{\Delta}$. The same is true for $\partial\Delta$. On the other hand, all points of $\Delta \setminus \{0\}$ together with 0 and the points of $\partial\Delta$ are limit points of $\Delta \setminus \{0\}$. Note that 0 and the points of $\partial\Delta$ are not in $\Delta \setminus \{0\}$.

(ii) The set $A = \{1/n : n \in \mathbb{N}\}$, where $\mathbb{N} = \{1, 2, 3, \ldots, n, \ldots\}$, has 0 as a limit point (regardless of whether the set is considered a subset of the line or the plane) and 0 is not in the set. Similarly, the set $A = \{e^{i\pi/n} : n \in \mathbb{N}\}$ has 1 as its only limit point, see Figure 2.3.

(iii) If A consists of the set of points that have both coordinates rational, then every point in the plane is a limit point of A. ●

A set is said to be *closed* if it contains all of its limit points. The union of a set A and its limit points is called the *closure* of A, and is denoted by $\overline{A}$. Some examples of closed sets in the plane are

(a) the empty set,
(b) the set of all complex numbers,

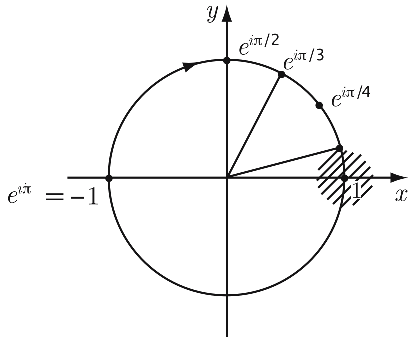


Figure 2.3. Description of limit point 1 of $\{e^{i\pi/n} : n \in \mathbb{N}\}$

(c) $\{z : |z| \geq r\}$, $r \geq 0$,
(d) $\{z : r_1 \leq |z| \leq r_2\}$, $0 \leq r_1 < r_2$,
(e) the union of any two closed sets,
(f) the intersections of any collection of closed sets,
(g) $\{z : |z| \leq 1\}$.

Some examples of sets that are not closed in the complex plane $\mathbb{C}$ are Δ, $\Delta \setminus \{0\}$, $\overline{\Delta} \setminus \{0\}$. Finally, we remark that the set $\overline{\Delta} \setminus \{0\}$ is neither closed nor open.

Theorem 2.3. *If z_0 is a limit point of A, then every neighborhood of z_0 contains infinitely many points of A.*

Proof. Assume that some deleted neighborhood of z_0 contains only a finite number of points of A. Let the points be $z_1, z_2, \dots, z_n$ and $\epsilon = \min_{i=1,2,\dots,n} |z_0 - z_i|$. Then $N'(z_0; \epsilon)$ contains no points of A, and z_0 can't be a limit point of A. ■

Corollary 2.4. *Every finite set is closed.*

Proof. The set contains all of its limit points—all "none" of them. ■

For the set $|z| \leq 1$, we would like to distinguish the interior points from the points on the unit circle. A point z_0 is called a *boundary point* of A if every neighborhood of z_0 contains points in A and points not in A (in the *complement* of A). The set of all boundary points of A is called *boundary* of A. For example, the circle $|z| = 1$ is the boundary for both the bounded set $|z| < 1$ and the unbounded set $|z| > 1$.

Remark 2.5. The boundary points determine the "openness" or "closedness" of a set. An open set cannot contain any of its boundary points, whereas a closed set must contain all of its boundary points (why?). Clearly, an interior point of a set A is a limit point of A but a limit point may or may not be an interior point of A. ●

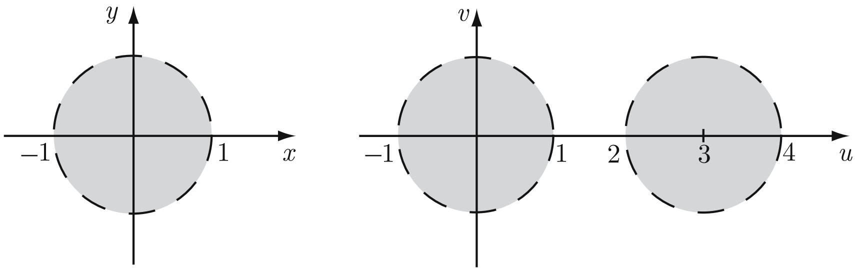


Figure 2.4. Description between open and connected sets

We would also like to distinguish between the two sets

$$A = \{z : |z| < 1\} \quad \text{and} \quad B = \{z : |z| < 1\} \cup \{z : |z-3| < 1\}.$$

Set A is "all one piece", while set B consists of two pieces (Figure 2.4). A set S is said to be *connected* if there do not exist disjoint open sets U and V satisfying the following conditions:

$$\text{(i) } U \cup V \supset S, \qquad \text{(ii) } U \cap S \neq \phi, \quad V \cap S \neq \phi.$$

In particular, if an open connected set can be expressed as the disjoint union of two open sets U and V, then either $U = \phi$ or $V = \phi$. Set A above is connected and set B is not.

An open connected set is called a *domain*. A *region* is a domain together with some, none, or all of its boundary points.[1] We might think that the counterpart of a real-valued function of a real variable being defined on an open set is a complex-valued function of a complex variable being defined on an open set. But this is not the case. Actually, the counterpart of an open interval in $\mathbb{R}$ is a domain. Note that an open interval in $\mathbb{R}$ is a connected subset of $\mathbb{R}$. Likewise a domain is open as well as connected. The "negative" definition for connectedness is sometimes difficult to visualize. But when the connected set is a domain, we have the following useful property.

Theorem 2.6. *Any two points in a domain can be joined by a polygonal line that lies in the domain.*

Proof. Choose a point z_0 in the domain D. It suffices to show that every point in D can be joined to z_0 by a polygonal line that lies in D. Let A denote the set of all points in D that can be so joined to z_0 and let B denote all those points that cannot. Note that $A \cup B = D$ and $A \cap B = \phi$. We wish to show that B is empty.

[1] The reader is warned that some authors use the term "region" for what we call a domain (following the modern terminology), and others make no distinction between the terms.

If a point z_1 is in A, then z_1 is in D. Since D is open, there exists an $\epsilon_1 > 0$ such that $N(z_1; \epsilon_1) \subset D$. But all the points in $N(z_1; \epsilon_1)$ can be joined to z_1 by a straight line segment. Therefore, each point in $N(z_1; \epsilon_1)$ must be in A, which means that A is an open set.

Similarly, if a point z_2 is in B, then there exists an $\epsilon_2 > 0$ such that $N(z_2; \epsilon_2) \subset D$. All the points in this neighborhood must also lie in B, for if some point $b \in N(z_2; \epsilon_2)$ could be joined to z_0 by a polygonal line, then the straight line segment from z_2 to b could be connected to the polygonal line from z_0 to b in order to form a polygonal line from z_0 to z_2. Thus B is an open set. Consequently, neither A nor B can contain any boundary points. Since D is connected, either A or B must be empty. But $z_0 \in A$, so that B is empty. This completes the proof. ■

Note that a domain may contain two points that cannot be joined by a single *straight* line segment, as is illustrated in Figure 2.5.

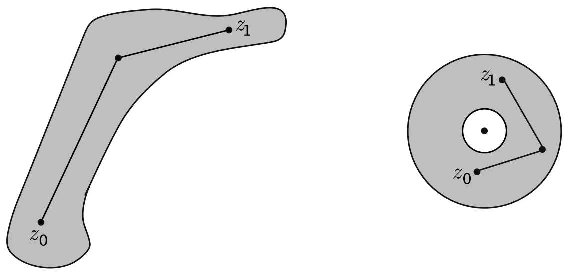


Figure 2.5. Connected domains

Remark 2.7. We could have required that the polygonal line of Theorem 2.6 be parallel to the coordinate axes. The only modification in the proof is the observation that any point in a disk can be joined to the center by combining a line segment parallel to the x axis with one parallel to the y axis. ●

The converse of Theorem 2.6 is also true: if any two points of an open set can be joined by a polygonal line, then the set is connected. The proof is left for the exercises. Also, in the exercises an example is given of a connected set, two of whose points cannot be joined by a polygonal line that lies in the set.

With the above definitions, we are furnished with a method for adequately characterizing most sets on either the line or the plane.

Examples 2.8. (i) Let $A = \{z \in \mathbb{C} : |z| \leq 1$, excluding the points $z_n = 1/n\,(n \in \mathbb{N})\}$. Then the set A is not open because the points on the unit circle have been included and is not closed because the limit points $z_n = 1/n\,(n \in \mathbb{N})$ have been excluded. The set is bounded, connected and has a boundary consisting of the unit circle, the points $z_n = 1/n$, and the origin.

(ii) Let $A = \{z \in \mathbb{C} : \operatorname{Re} z > 0\} \cup \{z : \operatorname{Re} z < -2\}$. This set is open, not closed, not bounded, and not connected. Its boundary consists of all points on the lines $\operatorname{Re} z = 0$ and $\operatorname{Re} z = -2$.

(iii) Let $A = \{z \in \mathbb{C} : -\pi/4 \leq \operatorname{Arg} z \leq \pi/4\}$. This set is connected, closed, not open, and not bounded. Its boundary consists of the origin together with the rays $\operatorname{Arg} z = \pi/4$ and $\operatorname{Arg} z = -\pi/4$. ●

Questions 2.9.

1. What alternative definitions of "bounded" might we have used?
2. What can we say about unions and intersections of open and closed sets?
3. What can we say about the complements of open and closed sets?
4. What sets are open (closed) in both the plane and the line?
5. What sets are both open and closed?
6. Can a set have infinitely many points without having a limit point?
7. What is the relation between the boundary points and limit points?
8. How does the closure of the intersection of two sets compare with the intersection of their closures?
9. What can you say about intersections and unions of connected sets?
10. What can you say about a set in which every pair of points can be joined by a straight line segment lying in the set?
11. How does the set described in the previous question compare to a set in which there exists a point that can be joined to any other point by a straight line segment lying in the set? What is an example of such a set?
12. What are the boundary points of a deleted neighborhood of z_0?
13. What are the boundary points of the complex plane?

Exercises 2.10.

1. Prove that a neighborhood of a point on the real line (an open interval) is an open set in $\mathbb{R}$.
2. Show that a set A of complex numbers is bounded if and only if, given $z_0 \in \mathbb{C}$, there exists a real number M such that $z \in N(z_0; M)$ for every $z \in A$. Can M be chosen independent of z_0?
3. Show that a set of complex numbers is bounded if and only if both the sets of its real and imaginary parts are bounded.
4. Describe the following sets.
   (a) $\{z \in \mathbb{C} : 1 < |z| < 2,$ excluding points for which $z \in \mathbb{R}\}$
   (b) $\{z \in \mathbb{C} : z = (x, y),\ x$ and y are rational$\}$
   (c) $\{x \in \mathbb{R} : x -$ irrational$\}$
   (d) $\{x \in \mathbb{R} : x \in \mathbb{Z}\}$
   (e) $\{n \in \mathbb{N} : \bigcup_{n=1}^{\infty}[1/n, n]\}$
   (f) $\{z \in \mathbb{C} : |z| > 2,\ |\operatorname{Arg} z| < \pi/6\}$
   (g) $\{z \in \mathbb{C} : |z + 1| < |z - i|\}$

(h) $\{z \in \mathbb{C} : |z+1| = |z-i|\}$
(i) $\{z \in \mathbb{C} : |\mathrm{Re}\, z| + |\mathrm{Im}\, z| = 1\}$.

5. Which of the following subsets are connected?
   (a) $D = \{z \in \mathbb{C} : |z| < 1\} \cup \{z \in \mathbb{C} : |z+2| \le 1\}$
   (b) $D = [0, 2) \cup \{2 + 1/n : n \in \mathbb{N}\}$.
6. Prove that the union of an arbitrary collection of open sets is open and that the intersection of a *finite* number of open sets is open. Also, show that $\cap_{n=1}^{\infty}\{z : |z| < 1/n\}$ is not an open set.
7. Show that a set is open if and only if its complement is closed.
8. Show that the intersection of an arbitrary collection of closed sets is closed and the union of a *finite* number of closed sets is closed.
9. Show that the limit points of a set form a closed set.
10. Show that $\overline{A}$, the closure of A, is the smallest closed set containing A.
11. Show that a set is connected if any two of its points can be joined by a polygonal line.
12. Show that if a set A is connected, then $\overline{A}$ is connected. Is the converse true?
13. Show that the union of two domains is a domain if and only if they have a point in common.

2.2 Sequences

A *sequence* $\{z_n\}$ of complex numbers is formed by assigning to each positive integer n a complex number z_n. The point z_n is called the nth term of the sequence. Care must be taken to distinguish between the terms of the sequence and the set whose elements are the term of the sequence. For example, the sequence $\{2, 2, 2, \dots\}$ has infinitely many terms (as do all sequences), but the set $\{2, 2, 2, \dots\}$ contains only one point. In general, when we discuss set-theoretic properties of a sequence, we will mean the set associated with the terms of the sequence.

A sequence $\{z_n\}$ is said to have a *limit* z_0 (*converge* to z_0), written

$$\lim_{n\to\infty} z_n = z_0 \quad \text{or} \quad z_n \to z_0,$$

if for every $\epsilon > 0$, there exists an integer N (depending on ϵ) such that $|z - z_0| < \epsilon$ whenever $n > N$. Geometrically, this means that every neighborhood of z_0 contains all but a finite number of terms of sequence (see Figure 2.6). We must point out that $z_n \to z_0$ is equivalent to $z_n - z_0 \to 0$. To illustrate, the sequence $\{1/n\}$ converges to 0; but the sequence $\{(-1)^n\}$, which oscillates between 1 and -1, does not converge. Examples of convergent sequences that appear frequently are

(a) $\displaystyle\lim_{n\to\infty} \frac{1}{n^p} = 0 \quad (p > 0)$
(b) $\displaystyle\lim_{n\to\infty} |z|^n = 0 \quad (|z| < 1)$
(c) $\displaystyle\lim_{n\to\infty} n^{1/n} = 1$.

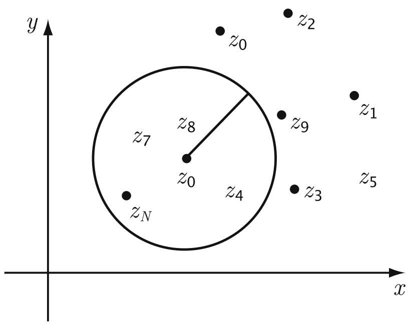


Figure 2.6. Geometric meaning of a convergence of a sequence

Example 2.11. We can easily see that $\{1+i^n\}_{n\geq 1}$ does not converge. Indeed, if $z_n = 1 + i^n$ then, for each fixed $k = 0, 1, 2, 3$,

$$z_{4n+k} = 1 + i^{4n+k} = 1 + i^k = \begin{cases} 2 & \text{if } k = 0 \\ 1+i & \text{if } k = 1 \\ 0 & \text{if } k = 2 \\ 1-i & \text{if } k = 3 \end{cases}$$

and so $\{1 + i^n\}$ diverges. Also we remark that $\{1 + i^n\}$ and $\{i^n\}$ diverge or converge together and so it suffices to deal with $\{i^n\}$ which is easier than the original sequence.

The convergence of the sequence $\{i^n/n\}_{n\geq 1}$ is easier to convince yourself of if you draw a figure representing these points. ●

There is a nice relationship between the convergence of a sequence of complex numbers and the convergence of its real and imaginary parts.

Theorem 2.12. *Let $z_n = x_n + iy_n$ be a sequence of complex numbers. Then $\{z_n\}$ converges to a complex number $z_0 = x_0+iy_0$ if and only if $\{x_n\}$ converges to x_0 and $\{y_n\}$ converges to y_0.*

Proof. The proof is simply a consequence of the inequalities

$$|\mathrm{Re}\, z|,\ |\mathrm{Im}\, z| \leq |z| \leq |\mathrm{Re}\, z| + |\mathrm{Im}\, z|.$$

To provide a detailed proof, we assume $\lim_{n\to\infty} x_n = x_0$ and $\lim_{n\to\infty} y_n = y_0$. Then given $\epsilon > 0$, there exist an integer N such that $n > N$ implies

$$|x_n - x_0| < \epsilon/2, \quad |y_n - y_0| < \epsilon/2. \tag{2.1}$$

From (2.1) we obtain

$$|z_n - z_0| = |x_n - x_0 + i(y_n - y_0)| \leq |x_n - x_0| + |y_n - y_0| < \epsilon,$$

and $\{z_n\}$ converges to z_0. Conversely, if we assume that $\lim_{n\to\infty} z_n = z_0$, the inequalities

$$|x_n - x_0| \le |z_n - z_0|, \quad |y_n - y_0| \le |z_n - z_0|$$

show that $\{x_n\}$ and $\{y_n\}$ converge to x_0 and y_0, respectively. ■

Theorem 2.12 essentially says that many properties of the complex sequences may be deduced from corresponding properties of real sequences. For example, the uniqueness of the limit of a complex sequence can be derived either directly or from the uniqueness property of real sequences.

A sequence of complex numbers $\{z_n\}$ is said to be *bounded* if there exists an $R > 0$ such that $|z_n| < R$ for all n. In other words, a sequence is said to be bounded if it is contained in some disk.

Since a convergent sequence eventually clusters about its limit, the next theorem is not too surprising.

Theorem 2.13. *A convergent sequence is bounded.*

Proof. If $\lim_{n\to\infty} z_n = z_0$, then $z_n \in N(z_0; 1)$ for $n > N$. Let

$$M = \max\{|z_1|, |z_2|, \dots, |z_N|\}.$$

Then, $|z_n| < M + |z_0| + 1$ for every n. ■

The converse of Theorem 2.13 is not true. The sequence $\{1, 2, 1, 2, \dots\}$ is bounded and not convergent, although the odd terms and even terms both form convergent sequences.

A *subsequence* of a sequence $\{z_n\}$ is a sequence $\{z_{n_k}\}$ whose terms are selected from the terms of the original sequence and arranged in the same order. For the sequence $z_n = (-1)^n$, we have subsequence $\{z_{2k}\}$ converging to 1 and subsequence $\{z_{2k-1}\}$ converging to -1.

The next theorem shows that a subsequence must be at least as "well behaved" as the original sequence.

Theorem 2.14. *If a sequence $\{z_n\}$ converges to z_0, then every subsequence $\{z_{n_k}\}$ also converges to z_0.*

Proof. Given $\epsilon > 0$, we have $z_n \in N(z_0; \epsilon)$ for $n > N$. Hence $z_{n_k} \in N(z_0; \epsilon)$ for $n_k > N$. Since $n_k \ge n$ (why?), and there can be at most N terms of the subsequence for which $|z_{n_k} - z_0| \ge \epsilon$. ■

We know that not all sets are bounded. However, if a set of real numbers is bounded, it has a "smallest" bound. A real number M is said to be the *least upper bound* (lub) of a nonempty set A of real numbers if

(i) $x \le M$ for every $x \in A$. That is A is bounded above by M and M is an upper bound for A.
(ii) For any $\epsilon > 0$, there exists a $y \in A$ such that $y > M - \epsilon$. That is, M is the smallest among all the upper bounds of A.

Similarly, the real number m is said to be the *greatest lower bound* (glb) of a nonempty set A if:

(i) $x \geq m$ for every $x \in A$; That is A is bounded below by m and m is a lower bound of A.

(ii) For any $\epsilon > 0$, there exists a $y \in A$ such that $y < m + \epsilon$. That is, m is the largest among all the lower bounds of A.

The *Dedekind property* states that *every nonempty bounded set of real numbers has a least upper bound and a greatest lower bound.* This is an amplified version of the result that $\mathbb{R}$ is complete. For a proof of this, see [R1].

As we have seen, the *converse* of Theorem 2.13 (even for real sequences) is not true. Bounded oscillating sequences need not converge. Eliminating the oscillation, however, will produce convergence. A real sequence $\{x_n\}$ is said to be *monotonically increasing* (*decreasing*) if $x_{n+1} \geq x_n (x_{n+1} \leq x_n)$ for every n. A sequence will be called *monotonic* if it is either monotonically increasing or monotonically decreasing.

Theorem 2.15. *Every bounded monotonic sequence of real numbers converges.*

Proof. Let the bounded sequence $\{x_n\}$ be monotonically increasing. According to the Dedekind property, there exists a least upper bound of $\{x_n\}$, call it x. By the definition of lub, given $\epsilon > 0$ there exists an integer N such that $x_N > x - \epsilon$. Since $\{x_n\}$ is monotonically increasing,

$$x - \epsilon < x_n \leq x \quad \text{for } n > N.$$

Hence $|x_n - x| < \epsilon$ for $n > N$, and $\{x_n\}$ converges to its least upper bound. The proof for monotonically decreasing sequences is identical, using the greatest lower bound instead of the least upper bound. ■

The examples we have seen of bounded sequences that did not converge did have convergent subsequences. To show that this is true in general, we need the following

Lemma 2.16. *Every sequence of real numbers contains a monotonic subsequence.*

Proof. Assume that the real sequence $\{x_n\}$ has the property that there are infinitely many n such that $x_k \leq x_n$ for every $k \geq n$. Let n_1 be the first such n with this property, n_2 the second, etc. Then $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$ is a monotonically decreasing subsequence of $\{x_n\}$.

On the other hand, if there are only finitely many n such that $x_k \leq x_n$ for every $k \geq n$, choose an integer m_1 such that no terms of the sequence $x_{m_1}, x_{m_1+1}, x_{m_1+2}, \ldots$ have this property. Let m_2 be the first integer greater than m_1 for which $x_{m_2} > x_{m_1}$. Continuing the process, we obtain a sequence $x_{m_1}, x_{m_2}, x_{m_3}, \ldots$ which is a monotonically increasing subsequence of $\{x_n\}$. This completes the proof. ■

Although the converse of Theorem 2.13 is not true, here is slightly a weaker version of it.

Theorem 2.17. *Every bounded sequence of complex numbers contains a convergent subsequence.*

Proof. Let $z_n = x_n + iy_n$, with $|z_n| \leq M$. Then $|x_n| \leq M$ and $|y_n| \leq M$. By Lemma 2.16, $\{x_n\}$ contains a monotonic subsequence $\{x_{n_k}\}$. By Theorem 2.15, $\{x_{n_k}\}$ converges.

Now consider the corresponding subsequence $\{y_{n_k}\}$ of $\{y_n\}$. This may not converge, but by Theorem 2.15, it does contain a convergent subsequence $\{y_{n_{k(l)}}\}$. By Theorem 2.14, $\{x_{n_{k(l)}}\}$ also converges. Applying Theorem 2.12, the sequence

$$z_{n_{k(l)}} = x_{n_{k(l)}} + iy_{n_{k(l)}}$$

is a convergent subsequence of $\{z_n\}$, and this completes the proof. ∎

What are the relationships between the limit of a sequence, the limit points of a sequence, and lub or glb of a sequence? The lub and glb are meaningless in the complex number system, although (as we have just seen) these notions for real numbers may be used to prove theorems about complex numbers. For the sequence $\{n/(n+1)\}$, 1 is the lub, the limit, and the unique limit point. If a convergent sequence has only finitely many distinct elements, it will have no limit points; however, we do have the following theorem.

Theorem 2.18. *A point z_0 is a limit point of a set A if and only if there is a sequence of distinct points in A converging to z_0.*

Proof. If a sequence $\{z_n\}$ of distinct points in A converges to z_0, then every neighborhood of z_0 contains all but a finite number (hence infinitely many) of points of $\{z_n\}$. Therefore, z_0 is a limit point of A.

To prove the converse, let z_0 be a limit point of A. For every integer n, choose a point $z_n \in N(z_0; 1/n) \cap A$. Since every neighborhood of A contains infinitely many distinct points, we may assume the points of the sequence $\{z_n\}$ to be distinct.

Given $\epsilon > 0$, choose N such that $1/N < \epsilon$. Then $z_n \in N(z_0; \epsilon)$ for $n > N$, and the sequence $\{z_n\}$ converges to z_0. ∎

Combining the previous two theorems, we obtain

Theorem 2.19. (Bolzano–Weierstrass) *Every bounded infinite set in the complex plane has a limit point.*

Proof. Choose any sequence of distinct points in the set. By Theorem 2.17, this sequence contains a convergent subsequence; and by Theorem 2.18, the limit of this convergent subsequence is a limit point of the set. This completes the proof. ∎

A sequence $\{z_n\}$ of complex numbers is said to be a *Cauchy sequence* if for every $\epsilon > 0$, there exists an integer N (depending on ϵ) such that

$$|z_m - z_n| < \epsilon \quad \text{whenever} \quad m, n > N.$$

What is the difference between a Cauchy sequence and a convergent sequence? Geometrically, for a convergent sequence, all but a finite number of points are close to a fixed point (the limit of the sequence), while for a Cauchy sequence all but a finite number of points are close to each other. We will show, for complex sequences, that these concepts are equivalent. Moreover, from our exercises, we see that the algebra of complex sequences is essentially the same as that for the real sequences studied in real-variable theory.

Theorem 2.20. (Cauchy Criterion) *The sequence $\{z_n\}$ converges if and only if $\{z_n\}$ is a Cauchy sequence.*

Proof. Assume $\{z_n\}$ converges to z_0. By the triangle inequality,

$$|z_m - z_n| = |z_m - z_0 + z_0 - z_n| \le |z_m - z_0| + |z_n - z_0|. \tag{2.2}$$

Given $\epsilon > 0$, both terms on the right side of (2.2) can be made less than $\epsilon/2$ for $m, n > N$. Hence $\{z_n\}$ is a Cauchy sequence.

Conversely, assume $\{z_n\}$ is a Cauchy sequence. Then for $n > N$, we have $|z_n - z_N| < 1$. That is,

$$|z_n| < |z_N| + 1 \quad \text{for } n > N.$$

Thus $\{z_n\}$ is a bounded sequence. By Theorem 2.17, $\{z_n\}$ contains a subsequence $\{z_{n_k}\}$ that converges to a point (say z_0).

We will show that $\{z_n\}$ also converges to z_0. Once again using the triangle inequality, we obtain

$$|z_n - z_0| = |z_n - z_{n_k} + z_{n_k} - z_0| \le |z_n - z_{n_k}| + |z_{n_k} - z_0|. \tag{2.3}$$

Given $\epsilon > 0$, there exists an integer N such that, for $n > N$,

$$\begin{cases} |z_n - z_{n_k}| < \epsilon/2 & \text{(because } \{z_n\} \text{ is Cauchy)}, \\ |z_{n_k} - z_0| < \epsilon/2 & \text{(because } \{z_{n_k}\} \text{ converges to } z_0). \end{cases} \tag{2.4}$$

Combining (2.3) and (2.4), we see that $|z_n - z_0| < \epsilon$ for $n > N$. Hence, $\{z_n\}$ converges to z_0, and the proof is complete. ■

Theorem 2.20 furnishes us with a general method for determining the convergence of a sequence of complex numbers even though we may not know in advance what its limit is. There are some systems in which not every Cauchy sequence converges. For instance, in the field of rational numbers, the Cauchy sequence $1, 1.41, 1.414, \ldots$ does not converge (because $\sqrt{2}$ is not rational). A system in which every Cauchy sequence converges is said to be *complete*. In Sprecher [S], it is shown that the real number system forms the only complete ordered field.

Example 2.21. Suppose that $z \neq 1$, but $|z| = 1$. Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} z^k = 0.$$

Indeed, as $(1-z)\sum_{k=1}^{n} z^k = z(1-z^n)$, we have

$$\left| \sum_{k=1}^{n} z^k \right| = \left| \frac{z(1-z^n)}{1-z} \right| \leq \frac{|z|(1+|z|^n)}{|1-z|} \leq \frac{2|z|}{|1-z|}$$

so that

$$\frac{1}{n}\left| \sum_{k=1}^{n} z^k \right| \leq \frac{2}{n}\left\{ \frac{|z|}{|1-z|} \right\} \to 0 \quad \text{as} \quad n \to \infty. \qquad \bullet$$

Questions 2.22.

1. When a sequence $\{z_n\}$ converges to z_0, is the limit z_0 unique?
2. Let $\{x_n\}$ and $\{y_n\}$ be real sequences. If $\{(x_n + y_n)\}$ converges, does this mean that both $\{x_n\}$ and $\{y_n\}$ converge? How does this question compare with Theorem 2.12?
3. How many subsequences are there for a given sequence?
4. Can unbounded sequences have limit points? What about monotonic unbounded sequences?
5. When will the least upper bound of a set be an element of the set?
6. Can a real sequence converge to a value other than lub or glb of the sequence?
7. Can a sequence have infinitely many limit points?
8. Can you think of a sequence that converges without knowing what its limit is?
9. How could Theorem 2.18 have been proved without appealing to Theorem 2.3?
10. What can be said of the sequence $b_n = \text{glb}\,\{a_n, a_{n+1}, a_{n+2}, \dots\}$, where $\{a_n\}$ is a real sequence? What if $\{a_n\}$ is bounded?
11. Suppose that $\{z_n\}$ converges. Does $\{|z_n|\}$ converge? Does $\{\arg z_n\}$ converge? Does $\{\text{Arg}\, z_n\}$ converge?
12. Suppose that both $\{\text{Arg}\, z_n\}$ and $\{|z_n|\}$ converge. Does $\{z_n\}$ converge?

Exercises 2.23.

1. Let $\{z_n\}$ converge to z_0 and w_n converge to w_0. Show that
   (a) $\lim_{n\to\infty} (z_n + w_n) = z_0 + w_0$,
   (b) $\lim_{n\to\infty} z_n w_n = z_0 w_0$,
   (c) $\lim_{n\to\infty} \frac{z_n}{w_n} = \frac{z_0}{w_0}$ provided $w_0 \neq 0$.

In particular, if

$$z_n = \frac{1+n+2i(n-1)}{n} \quad \text{and} \quad w_n = \frac{n^{1/2}+2i(3+4n^3)}{n^3},$$

find z_0, w_0 and z_0/w_0.

2. Show that no sequence having more than one limit point can converge.
3. If $\{z_n\}$ converges, show that $\{|z_n|\}$ converges. Is the converse true?
4. Which of the following sequences are convergent?
   (a) $\{i^n\}$
   (b) $\{z_0^n\}$, where $|z_0| < 1$
   (c) $\left\{\dfrac{\cos n + i \sin n}{n}\right\}$
   (d) $\left\{\dfrac{1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}}{n}\right\}$
   (f) $\left\{e^{n\pi i/3} + \left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right)^n\right\}$
   (g) $\left\{e^{n\pi i/6} + \left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)^n\right\}$
   (f) $\left\{\frac{n\cos(n\pi)}{2n+1}\right\}$
   (i) $\left\{\sin\left(\frac{n\pi}{8}\right)\right\}$.
5. If $\{z_n\}_{n\geq 1}$ converges to 0, prove that $\{\frac{1}{n}\sum_{k=1}^n z_k\}$ converges to 0. Then show that $\{z_n\}$ converging to z_0 implies that $\{\frac{1}{n}\sum_{k=1}^n z_k\}$ converges to z_0.
6. Give an example of a sequence that
   (a) does not converge, but has exactly one limit point;
   (b) has n limit points, for any given integer n;
   (c) has infinitely many limit points.
7. Prove that the subsequential limits (the limits of all possible subsequences) of a sequence $\{z_n\}$ form a closed set.
8. Let $\{z_n\}$ be a sequence having the following property: Given $\epsilon > 0$, there exists an integer N such that for $n > N$, $|z_{n+1} - z_n| < \epsilon$. Give an example to show that $\{z_n\}$ need not be a Cauchy sequence.
9. Let $s_n = \sum_{k=1}^n 1/k!$. Use the Cauchy criterion to show that $\{s_n\}$ converges.

2.3 Compactness

The union of the open intervals $(n - \frac{1}{2}, n + \frac{1}{2})$ for $n = 1, 2, 3, \ldots$ contains the set of positive integers. Each interval is important in that the removal of any one of them will leave a positive integer uncovered. For the bounded set $S = \{x \in \mathbb{R} : 0 < x < 1\}$, the union of open intervals $(1/n, 1)$ for $n = 2, 3, 4, \ldots$ contains S. While the removal of any one of these intervals will prevent the union of the remaining intervals from covering S, the set S is not contained in any finite subcollection of the intervals.

A set is said to be *countable* if its elements can be put in a one-to-one correspondence with a subset of positive integers. A collection $\{O_\alpha\}$ of open sets is called an *open cover* of a set S if $S \subset \bigcup_\alpha O_\alpha$. Note that the collection $\{O_\alpha\}$ may contain uncountably many sets. A set S is *compact* if every open cover of S contains a finite subcover.

We have seen that neither the set of positive integers nor the open interval $(0, 1)$ is compact. However, any finite set is compact because for any open cover we have a finite subcover formed by associating with each point one of the open sets containing the point.

The definition of compactness is not always easy to apply. We would like to work with a more geometrically intuitive method for determining compactness. To this end we will need the following.

Lemma 2.24. *Let $\{I_n\}$ be a sequence of closed and bounded intervals on the real line. If $I_{n+1} \subset I_n$ for every n and the length of I_n approaches 0 as $n \to \infty$, then there is exactly one point in common to all I_n.*

Proof. Let $I_n = \{x : a_n \le x \le b_n\}$. By hypothesis,

$$a_n \le a_{n+1}, \quad b_{n+1} \le b_n \quad (n = 1, 2, 3, \dots)$$

and

$$\lim_{n\to\infty} (b_n - a_n) = 0. \tag{2.5}$$

The sequences $\{a_n\}$ and $\{b_n\}$ are both monotonic and bounded $(a_n, b_n) \in [a_1, b_1]$ for every n. By Theorem 2.15 both sequences must converge; and by (2.5), they must converge to the same point, call it x_0. Since $x_0 = \operatorname{lub}\{a_n\} = \operatorname{glb}\{b_n\}$,

$$x_0 \in [a_n, b_n] \quad \text{for every } n$$

(see Figure 2.7). There cannot be another point x_1 in all the I_n. For, if x_1 ($\ne x_0$) were less than (resp. greater than) x_0, then x_0 would not be the lub $\{a_n\}$ (resp. glb $\{b_n\}$). ∎

Note that Lemma 2.24 is not true if *closed* is replaced by *open*. The collection of intervals $\{(0, 1/n) : n \in \mathbb{N}\}$ satisfy the hypotheses, although $\bigcap_{n=1}^\infty (0, 1/n) = \phi$.

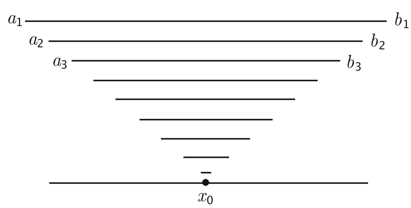


Figure 2.7.

Lemma 2.25. *Let $\{S_n\}$ be a sequence of closed and bounded rectangles in the plane. If $S_{n+1} \subset S_n$ for every n and the length of the sides of S_n approaches 0 as $n \to \infty$, then there is exactly one point in common to all the S_n.*

Proof. Let $\{I_n\}$ and $\{J_n\}$ be the projections of $\{S_n\}$ into real and imaginary axes respectively. Then $\{I_n\}$ and $\{J_n\}$ satisfy the conditions of Lemma 2.24. If

$$\{x_0\} = \cap_{n=1}^{\infty} I_n \quad \text{and} \quad \{y_0\} = \cap_{n=1}^{\infty} J_n,$$

then (see Figure 2.8) $\{z_0\} = \{(x_0, y_0)\} = \cap_{n=1}^{\infty} S_n$. ■

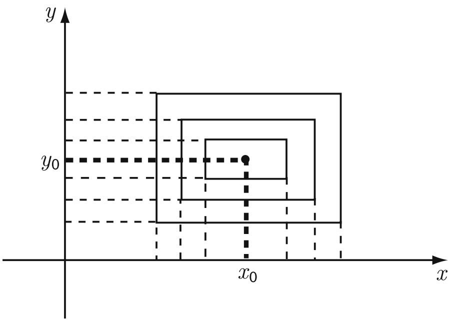


Figure 2.8.

Theorem 2.26. (Heine–Borel) *Every closed and bounded set is compact.*

Proof. Let S be a closed and bounded set. Assume $\{O_\alpha\}$ is an open cover of S that has no finite subcover. Since S is bounded, it is contained in some square S_0 whose vertices are $z = \pm a \pm ai$. The coordinate axes divide S_0 into four equal subsquares. At least one of these squares (call it S_1) has the property that $S \cap S_1$ cannot be covered by a finite subfamily of $\{O_\alpha\}$ (why?). We now divide S_1 into four *more* equal closed subsquares (see Figure 2.9). Again, for at least one of these squares, denoted by S_2, there is no finite subfamily of $\{O_\alpha\}$ that covers $S \cap S_2$.

We can continue the process indefinitely, forming a sequence $\{S_n\}$ of closed squares for which there is no finite subfamily of $\{O_\alpha\}$ that covers $S \cap S_n$. Note that the length of any side of S_n is $a/(2^{n-1})$. By Lemma 2.25, there is exactly one point, denoted by z_0, common to all squares S_n. This point z_0 must be a limit point of S, and hence an element of S.

Let O_{α_0} be an element of the cover $\{O_\alpha\}$ that contains z_0. Since O_{α_0} is an open set, $N(z_0; \epsilon) \subset O_{\alpha_0}$ for some $\epsilon > 0$. But $S_n \subset N(z_0; \epsilon)$ for n sufficiently large. Thus $S_n \cap S$ has a finite subcover of $\{O_\alpha\}$, namely, one element: O_{α_0}. This contradiction, concludes the proof. ■

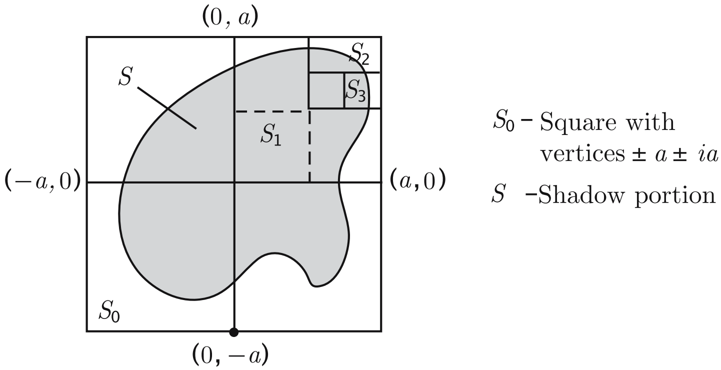


Figure 2.9. Illustration for Heine–Borel theorem

The gist of the above argument is that if no finite subcollection of $\{O_\alpha\}$ covers S, then no finite subcollection covers a carefully chosen sequence of subsets of S. On the other hand, this sequence of subsets can be made small enough to be contained in one of the open sets of the cover.

We are now ready to collect some of the important results of the last two sections to obtain the following major theorem.

Theorem 2.27. *Let S be a subset of the complex plane $\mathbb{C}$. The following statements are equivalent:*

(i) *S is closed and bounded.*
(ii) *S is compact.*
(iii) *Every infinite subset of S has a limit point in S.*
(iv) *Every sequence in S has a subsequence that converges to a point in S.*

Proof. The Heine–Borel theorem states that (i) implies (ii). We will show that (ii) implies (iii), (iii) implies (iv), and (iv) implies (i). Since each statement is clearly correct if S is a finite set, we may suppose that S is infinite.

Assume that (ii) holds. If A is an infinite subset of S having no limit point in S, then for every point in $S \setminus A$ we can find a neighborhood containing no points of A. Furthermore, for every point in A we can find a neighborhood containing no other points of A. The collection of all such neighborhoods is an open cover of S for which there is no finite subcover, contradicting the compactness of S.

Assume (iii) holds. Let $\{z_n\}$ be a sequence of distinct points in S. (Why is it sufficient to consider only such sequences?) By hypothesis, there exists a limit point z_0 of $\{z_n\}$ with $z_0 \in S$. By Theorem 2.18, there is some subsequence of $\{z_n\}$ converging to z_0.

Assume (iv) holds. If S is unbounded, then there exists a sequence of points $\{z_n\}$ in S such that $|z_n| > n$ for every n. Let $\{z_{n_k}\}$ be an arbitrary subsequence of $\{z_n\}$. For any point $z_0 \in S$, $N(z_0; 1)$ can contain no points of $\{z_{n_k}\}$ for which $n_k > |z_0| + 1$. Hence $\{z_{n_k}\}$ cannot converge to any point in

S, contradicting our assumption: To show that S is closed, let z_0 be a limit point of S. By Theorem 2.18, there is sequence of distinct points $\{z_n\}$ of S converges to z_0. By Theorem 2.14, every subsequence of $\{z_n\}$ converges to z_0. According to (iv), z_0 must therefore be in S. This completes the proof. ■

Compactness is a nice property because reducing an open cover to a finite subcover often means that only a finite number of points need be considered in proving that a set has a certain property. For this reason, when we have compactness, many local properties (properties that hold in a neighborhood of each point in a set) can be shown to be global or uniform (a property of the set as a whole).

For example, from the fact that each point may be covered by a bounded neighborhood, we deduced that a compact set is bounded. Also, if each point in a compact set is a positive distance from a fixed point, the set itself is a positive distance from the point (see Exercise 2.29(3)). This, of course, is not true in general. Each point of the open interval $(0,1)$ is a positive distance from 0, but we can not find a positive real number between 0 and the set.

What makes the addition of one or two points so important? Let us compare the open interval $(0,1)$ with the closed interval $[0,1]$. As we saw earlier, $\bigcup_{n=2}^{\infty}(1/n,1)$ is an open cover of $(0,1)$, that has no finite subcover. This cover does not contain the points $\{0\}$ and $\{1\}$. If these points were added to the set, intervals like $(-\epsilon,\epsilon)$ and $(1-\epsilon,1+\epsilon)$ would also have to be added to obtain a cover. But then $(-\epsilon,\epsilon)$, $(1-\epsilon,1+\epsilon)$ and $\bigcup_{n=2}^{N}(1/n,1)$ for $N>1/\epsilon$ would be a finite subcover.

Questions 2.28.

1. What can we say about the finite union (intersection) of compact sets?
2. What can we say about the infinite union (intersection) of compact sets?
3. What can we say about the complement of a compact set?
4. What can we say about Cauchy sequences in compact sets?
5. When can we say that every subset of a compact set is compact?
6. We have seen that the removal of one point from a set may destroy the compactness. How many points may be *added* to a set to destroy compactness?
7. What kind of generalizations to Lemma 2.25 might we have for compact sets?
8. Can we talk about “infinity” being a limit point?

Exercises 2.29.

1. Show that the union of any bounded set and its limit points is a compact set.
2. Show that a compact set of real numbers contains its greatest lower bound and its least upper bound. Can this occur for a set of real numbers that is not compact?

3. If S is compact and $z_0 \notin S$, prove that $\operatorname{glb}_{z\in S}|z-z_0| > 0$.
4. If $\{S_n\}$ is a sequence of nonempty compact sets with $S_{n+1} \subset S_n$ for every n, show that $\bigcap_{n=1}^{\infty} S_n \neq \phi$.
5. In Theorem 2.27, prove as many different implications as you can.
6. Show that the set of rational numbers are countable.
7. Show that any open cover of a subset of the plane has a *countable* subcover.

2.4 Stereographic Projection

Thus far, infinite limits have been carefully avoided. Consider the three real sequences:

$$a_n = n, \quad b_n = \begin{cases} n & \text{if } n \text{ is odd} \\ 1 & \text{if } n \text{ is even} \end{cases}, \quad c_n = (-1)^n n.$$

Even though all three sequences grow arbitrarily large, we do not want to say they all approach infinity. From our knowledge of finite limits, it seems appropriate that $\{a_n\}$ should approach infinity and that $\{b_n\}$ should not, since a subsequence of $\{b_n\}$ converges to 1. A case for $\{c_n\}$ can be made either way. The standard approach is to introduce the symbols $\pm\infty$, and adjoin them to the real numbers. The set $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty, -\infty\}$ is known as the *extended real number system.* In the extended real number system, we use the following conventions:

$$\begin{cases} \pm\infty + a = \pm\infty = a \pm \infty & \text{for } a \in \mathbb{R} \\ \infty \cdot a = a \cdot \infty = \infty & \text{for } a \in \mathbb{R}_\infty \backslash \{0\} \\ \dfrac{a}{\infty} = 0 & \text{for } a \in \mathbb{R}\backslash\{0\} \\ \dfrac{a}{0} = \infty & \text{for } a \in \mathbb{R}_\infty \backslash \{0\}. \end{cases}$$

The expressions $\infty + \infty = \infty$, $-\infty - \infty = -\infty$ hold while $\infty - \infty$ is not defined. In the extended real number system, $\{c_n\}$ does not converge because $\{c_{2n}\}$ approaches $+\infty$ and $\{c_{2n+1}\}$ approaches to $-\infty$.

A perfectly logical, if somewhat unusual, approach is to adjoin only one point, ∞, to $\mathbb{R}$. We then say that a sequence $\{a_n\}$ approaches ∞, written $\lim_{n\to\infty} a_n = \infty$, if, for any preassigned real number M, all but a finite number of terms lie outside the interval $(-M, M)$. According to this definition, the sequence $\{(-1)^n n\}$ does approach ∞.

This latter approach can be thought to arise from the former by grabbing the two points $-\infty$ and $+\infty$ (with two very long arms) and bringing them together. The real number line is then transformed into a circle. We now make this geometric notion more precise. Consider the unit circle $x^2 + y^2 = 1$. For any real number a, draw the straight line joining the points $(a, 0)$ and $(0, 1)$. This line intersects the unit circle at $(0, 1)$ and one other point (x_1, y_1), which we identify with the real number a. For example, points in the open interval

$(-1, 1)$ are identified with points in the lower half of the circle, the points -1 and 1 are identified with themselves, and the points outside the interval $[-1, 1]$ are identified with points in the upper half of the circle (see Figure 2.10).

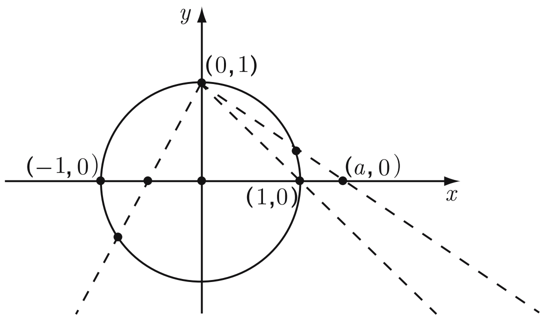


Figure 2.10. Illustration for the existence of $+\infty$ and $-\infty$ in $\mathbb{R}$

Observe that points close to one another on the real line are always identified with points close to one another on the circle. The converse is not true. Points "far out" in the positive and negative directions are identified with points close to one another on the unit circle. In fact, the greater the absolute value of a real number the closer is its identification with a point near $(0, 1)$, the only point on the unit circle not identified with a real number. For this reason, we identify the point $(0, 1)$ with the point ∞. This provide us with a one-to-one correspondence between points in the set $\mathbb{R} \cup \{\infty\}$ and the points on the unit circle. Since the set of real numbers is not compact, the identification of $\mathbb{R} \cup \{\infty\}$ with (compact) circle is called a *one-point compactification* of the real numbers.

Was the elimination of $-\infty$ worth all this effort? Not really. In fact, it is actually useful for $-\infty$ to mean "less than any real number". The set $\mathbb{R} \cup \{\infty\}$ was introduced in order to properly motivate our study of the extended complex plane. Consider the complex sequence $\{z_n\}$ defined by $z_n = n(\cos\theta + i\sin\theta)$, where $0 \le \theta \le 2\pi$. For each different value of θ, $\{z_n\}$ approaches ∞ along a different ray. Furthermore, since the complex numbers are not ordered, the symbol $-\infty$ would have no more meaning than the symbol $i\infty$.

In the case of complex numbers, by an *M neighborhood of* ∞, denoted by $N(\infty; M)$, we mean the set of all points whose absolute value is greater than M. That is the exterior of the disk with radius M and center at the origin. The sequence $\{z_n\}$ is said to approach ∞ if for any $M > 0$, $z_n \in N(\infty; M)$ for all but a finite number of n.

If we adjoin the point at ∞ to the set of complex numbers, we obtain the *extended complex number system*. Sometimes $\mathbb{C}$ is referred to as the finite complex plane and is designated also by $|z| < \infty$. Then $\mathbb{C} \cup \{\infty\} := \mathbb{C}_\infty$ is called the *extended complex plane*. Note that the extended complex number system is conceptually different from the extended real number system, in which two points ($+\infty$ and $-\infty$) are added. We first make the following algebraic rules as definitions:

$$\begin{cases} \infty \pm z = \infty = z \pm \infty \quad \text{for } z \in \mathbb{C} \\ \infty \cdot z = z \cdot \infty = \infty \quad \text{for } z \in \mathbb{C}_\infty \setminus \{0\} \\ \dfrac{z}{\infty} = 0 \quad \text{for } z \in \mathbb{C} \setminus \{0\} \\ \dfrac{z}{0} = \infty \quad \text{for } z \in \mathbb{C}_\infty \setminus \{0\}. \end{cases}$$

There is a difficulty in assigning meaning to the expressions $\infty + \infty$, $\infty - \infty$ ∞/∞, $\infty \cdot 0$ and $0/0$ and so none of these expressions has meaning in $\mathbb{C}_\infty$. The one-point compactification, $\mathbb{C}_\infty := \mathbb{C} \cup \{\infty\}$, of the plane has geometric model similar to that of the one-point compactification of the line, with the unit circle being replaced by the unit sphere

$$S = \{(x, y, z) : x^2 + y^2 + u^2 = 1\}$$

in the 3-dimensional Euclidean sphere in $\mathbb{R}^3$.

Identify the complex number $a + ib$ with the point $(a, b, 0)$ in $\mathbb{R}^3$. By doing so, we are free to imagine $\mathbb{C}$ as an object sitting inside $\mathbb{R}^3$ as xy plane. Having made this identification, for every number $a+ib$ in the complex plane, draw the straight line in $\mathbb{R}^3$ connecting the points $(a, b, 0)$ and $(0, 0, 1)$. This line intersects the sphere $x^2 + y^2 + u^2 = 1$ at $(0, 0, 1)$ and at one other point (x_1, y_1, u_1). The projection from the point $(0, 0, 1)$ on the sphere to the point $(a, b, 0)$ in the complex plane is called a *stereographic projection* (see Figure 2.11). The sphere S is called the Riemann sphere.

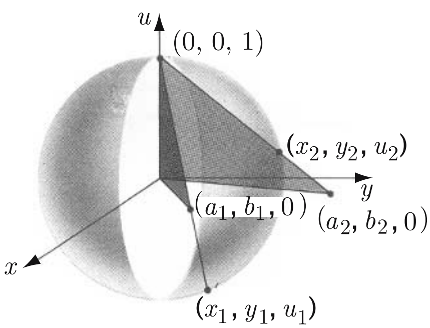


Figure 2.11. Stereographic projection

This one-to-one correspondence covers all points in the finite complex plane and all points in the sphere except $(0, 0, 1)$. The point at ∞ in the extended complex- number system is identified with the point $(0, 0, 1)$, sometimes called the *north pole*. Note that a neighborhood of ∞ in the complex plane corresponds to the interior of an arctic circle whose center is the north pole.

To find specifically the point (x_1, y_1, u_1) on the sphere identified with the point $(a, b, 0)$, observe that the three points

$$(0,0,1), \quad (x_1, y_1, u_1), \quad \text{and} \quad (a, b, 0)$$

are collinear. Hence,

$$\frac{x_1 - 0}{a} = \frac{y_1 - 0}{b} = \frac{u_1 - 1}{-1} = t \tag{2.6}$$

for some real scaler t. But

$$x_1^2 + y_1^2 + u_1^2 = (at)^2 + (bt)^2 + (1-t)^2 = 1, \quad \text{i.e.,} \quad (a^2 + b^2 + 1)t^2 = 2t.$$

Solving for t, we obtain

$$t = \frac{2}{a^2 + b^2 + 1} = 1 - u_1$$

as $t = 0$ corresponds to $(0, 0, 1)$, the north pole. In view of (2.6), the complex number $a + ib$ is then identified with the point

$$(x_1, y_1, u_1) = \left(\frac{2a}{a^2 + b^2 + 1}, \frac{2b}{a^2 + b^2 + 1}, \frac{a^2 + b^2 - 1}{a^2 + b^2 + 1} \right). \tag{2.7}$$

Rewriting (2.7), we identify the complex number $z = x + iy$ with the point on the sphere

$$\left(\frac{2x}{|z|^2 + 1}, \frac{2y}{|z|^2 + 1}, \frac{|z|^2 - 1}{|z|^2 + 1} \right).$$

From the second formula for t and (2.6), we conclude that

$$a = \frac{x_1}{1 - u_1} \quad \text{and} \quad b = \frac{y_1}{1 - u_1}.$$

Consequently, we identify the point (x, y, u) in $S \setminus \{(0,0,1)\}$ with the complex number in the plane

$$\left(\frac{x}{1-u} \right) + i \left(\frac{y}{1-u} \right).$$

For instance the points, $z = 0$ and $z = 1 - i$ correspond to the points $(0, 0, -1)$ and $(2/3, -2/3, 1/3)$, respectively.

Questions 2.30.

1. Which theorems for finite limit remain true for infinite limits?
2. What is the relationship between unbounded sets and neighborhoods of ∞?

3. How might we define ∞ to be a limit point of a sequence?
4. What might the symbol "$i\infty$" mean?
5. What might the symbol "$-i\infty$" mean?
6. What happens to the points on the unit circle in the complex plane under stereographic projection?
7. Could we have identified the complex plane with a different sphere?
8. What would be a one-point compactification of $\mathbb{R}^n$?
9. How are the images on the Riemann sphere of z and $\overline{z}$ related?
10. How are the images on the Riemann sphere of z and $-z$ related? How about for z and $-\overline{z}$?
11. What is the image of the line $x+y=1$ in the complex plane, on the Riemann sphere?

Exercises 2.31.

1. Show that a sequence having a finite limit point cannot approach ∞.
2. If $\{z_n\}$ approaches ∞ and $\{w_n\}$ is bounded, show that $\{(z_n+w_n)\}$ approaches ∞.
3. Show that $\{z_n\}$ approaches ∞ if and only if $\{|z_n|\}$ approaches ∞.
4. Given a point (x_1, y_1, u_1) on the unit sphere, find its corresponding point in the complex plane.
5. Show that a circle on the sphere that does not pass through the north pole corresponds to a circle in the complex plane.
6. Show that a circle on the sphere passing through the north pole corresponds to a straight line in the complex plane.
7. Show that we may identify, by stereographic projection, the complex plane with the sphere $x^2+y^2+(u-\frac{1}{2})^2=(\frac{1}{2})^2$.
8. Consider two antipodal points (x,y,u) and $(-x,-y,-u)$ on the Riemann sphere. Show that their stereographic projections z and z' are related by $zz'=-1$. Give a geometric interpretation.
9. Show that the image of the circle $|z|=\sqrt{3}$ under the stereographic projection is the set of all points (x_1,y_1,u_1) in the sphere described by $x_1^2+y_1^2=3/4$ and $u_1=1/2$.

2.5 Continuity

A (single-valued) *function* or mapping f from a set A into a set B, written $f:A\to B$, is a rule that associates with each element x of A a unique element $f(x)$, the value of f at x, of B. The set A is called the *preimage* (or the domain set) of f and the subset of B associated with the element of A is called the *image* of f and is denoted by $f(A)$, i.e. $f(A)=\{f(x) : x\in A\}$. If the set B, called the *range* of the function, is equal to $f(A)$, the function is said to be *onto*. If no two elements of A are mapped onto the same element in B, the function is said to be *one-to-one* on A. By $f(a)=b$, we will mean that the element $a\in A$ is mapped onto the element of $b\in B$.

For each $b \in B$, we define $f^{-1}(b)$ to be the set of elements in A whose image is b. Note that $f^{-1}(b)$ may be empty if f is not onto. However, if f is one-to-one and onto, $f^{-1} : B \to A$ is also a one-to-one and onto function, called the inverse function of f.

Example 2.32. The function $w = f(z) = az + b$, $a \neq 0$, is one-to-one in $\mathbb{C}$ and the inverse function is defined by $z = (w - b)/a$. Note that both are defined in the whole plane $\mathbb{C}$.

On the other hand, the function f defined by $f(z) = z + 3z^2$ is not one-to-one in $|z| < 1$. For,

$$z_1 + 3z_1^2 = z_2 + 3z_2^2 \implies (z_1 - z_2) = 3(z_2 - z_1)(z_1 + z_2)$$

which implies $(z_1 - z_2)[1 + 3(z_1 + z_2)] = 0$ and we see that the last equality is true when $z_1 + z_2 = -1/3$. But there are many points $z_1,\ z_2 \in \varDelta$ such that $z_1 + z_2 = -1/3$. However, this function is one-to-one in $|z| < 1/6$. ●

We have tacitly been dealing with functions. For example, a sequence of real numbers is a function $f : \mathbb{N} \to \mathbb{R}$ and a sequence of complex numbers is a function $f : \mathbb{N} \to \mathbb{C}$, where $\mathbb{N}$ is the set of positive integers. In stereographic projection, a one-to-one function was found that mapped the extended complex plane onto the unit sphere. The reader (hopefully) is familiar with some of the properties of real-valued functions of a real variable, i.e., functions mapping sets of real numbers onto sets of real numbers. For example, the function $y = f(x) = x^2$, mapping the real variable x onto the real variable x^2, takes the set of real numbers onto the set of nonnegative real numbers, the closed interval $[0, 1]$ onto itself, and so on.

Remark 2.33. Strictly speaking, f stands for the function and $f(x)$ for the value of the function at the point x. However, when there is no ambiguity, we will sometimes use the time-honored notational abuse of referring to $f(x)$ as a function.

For $z = x + iy$, the complex-valued function $f(z)$ can be viewed as a function of the complex variable z or as a function of two real variables x and y. ●

For example, the function $f(z) = z^2$ may be expressed as

$$w = f(z) = f(x, y) = (x + iy)^2 = x^2 - y^2 + i(2xy),$$

where

$$\operatorname{Re} f(z) = u(x, y) = x^2 - y^2 \quad \text{and} \quad \operatorname{Im} f(z) = v(x, y) = 2xy.$$

For this function, the points $(2, 1)$, $(1, 2)$, and $(3, -1)$ are mapped onto the points $(3, 4)$, $(-3, 4)$, and $(8, -6)$ respectively.

Just as a real-valued function of a real variable may be viewed as a mapping from the x axis to the y axis, so may a complex-valued function of a complex

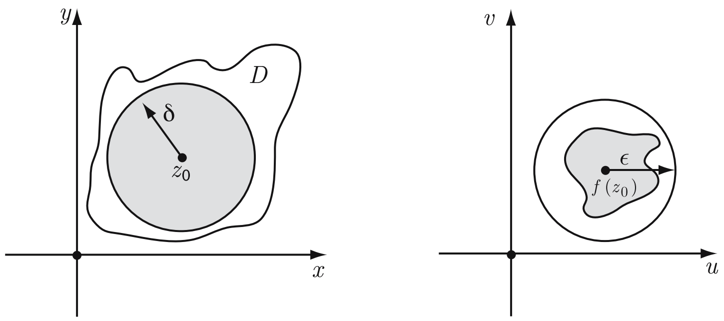


Figure 2.12. Concept of continuity at z_0

variable be viewed as a mapping from the xy plane (z plane) to the uv plane (w plane). While the y axis may be placed vertically on the x axis to obtain a complete two-dimensional picture of the real-valued function $y = f(x)$, the z plane and w plane must stay apart, at least in this three-dimensional world. In this book, we mostly deal with functions $f : A \to \mathbb{C}$ where A is a subset of $\mathbb{C}$.

In Chapter 3, we will be concerned with functions that map certain regions in the z plane onto certain regions in the w plane. Right now we have the more modest task of determining a class of functions that map points near one another in the z plane onto points near one another in the w plane.

A function $f(z)$, defined in a domain D, is said to be *continuous* at a point $z_0 \in D$ if for every $\epsilon > 0$, there exists a $\delta > 0$ (δ depending on ϵ and z_0) such that

$$|f(z) - f(z_0)| < \epsilon, \quad \text{whenever} \quad |z - z_0| < \delta. \tag{2.8}$$

Geometrically, this means that, for every neighborhood of $f(z_0)$ in the w plane, there corresponds a neighborhood of z_0 in the z plane whose image is contained in the neighborhood of $f(z_0)$. More formally, for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$f(N(z_0;\delta)) \subset N(f(z_0);\epsilon) \tag{2.9}$$

(see Figure 2.12). If a function is continuous at every point of D, the function is said to be continuous in the domain D. A function $f : A \to \mathbb{C}$ is discontinuous (or has a discontinuity) at z_0 if $z_0 \in A$, yet f is not continuous at $z = z_0$.

Remark 2.34. We will use (2.8) and (2.9) interchangeably. The reader should convince himself of their equivalence and strive to be equally proficient with both.

Also, we will have occasion to discuss the continuity of a function in a region R that includes boundary points. By an ϵ neighborhood of a boundary point $z_0 \in R$, we will mean $N(z_0;\epsilon) \cap R$, and will call this an open set relative

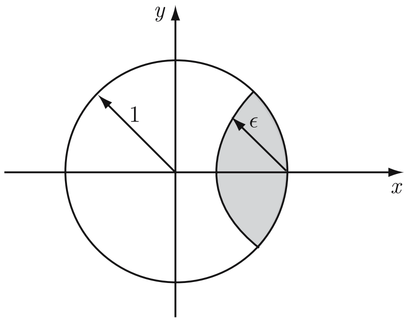


Figure 2.13. ϵ-neighborhood of a boundary point

to the region R. See Figure 2.13 for an ϵ neighborhood of a boundary point of the closed unit disk $|z| \le 1$. ●

If $f(z)$ is continuous at z_0, we write $\lim_{z\to z_0} f(z) = f(z_0)$. A function may have a limit at a point without being continuous at the point. We say that $\lim_{z\to z_0} f(z) = L$ if for every neighborhood of L, there is a *deleted* neighborhood of z_0 whose image is contained in the neighborhood of L. If $L = f(z_0)$, the function is continuous at z_0 and the word "deleted" may be deleted from this definition.

Examples 2.35. (i) Let

$$f(z) = \begin{cases} z^2 & \text{if } z \neq 2, \\ 5 & \text{if } z = 2. \end{cases}$$

For this function, $\lim_{z\to 2} f(z) = 4$ although the function is not continuous at $z = 2$.

(ii) Let

$$f(z) = \begin{cases} \dfrac{z-2}{z^2-4} & \text{if } z \neq 2, \\ 4 & \text{if } z = 2. \end{cases}$$

Then $\lim_{z\to 2} f(z) = \lim_{z\to 2} 1/(z+2) = 1/4 = L$. Here $L \neq f(2)$. Hence f has a limit as $z \to 2$ but is not continuous at $z = 2$.

(iii) If $\lim_{z\to a} f(z) = L$, then for a given $\epsilon > 0$ there exists $\delta > 0$ such that

$$|\,|f(z)| - |L|\,| \le |f(z) - L| < \epsilon \quad \text{whenever} \quad 0 < |z-a| < \delta$$

and therefore,

$$\lim_{z\to a} |f(z)| = |L|.$$

Clearly, if $L = 0$, $\lim_{z\to a} |f(z)| = |L|$ iff $\lim_{z\to a} f(z) = L$. What happens if $L \neq 0$? More precisely, if $\lim_{z\to a} |f(z)| = L'$, then is it always the case that $\lim_{z\to a} f(z)$ exists? Remember that if $\lim_{z\to a} f(z) = L$, then $|L| = L'$ and therefore, we have to examine when equality holds in

$$||f(z)| - L'| = ||f(z)| - |L|| \le |f(z) - L|.$$

Equality would imply that

$$\mathrm{Re}\,(f(z)\overline{L}) = |f(z)|\,|L| \quad \text{or} \quad |f(z)| = \mathrm{Re}\left(f(z)\frac{\overline{L}}{|L|}\right) = \mathrm{Re}\,(e^{i\theta}f(z))$$

where $\theta = \mathrm{Arg}\,(\overline{L}/|L|)$, or equivalently,

$$|e^{i\theta}f(z)| = \mathrm{Re}\,(e^{i\theta}f(z))$$

so that $e^{i\theta}f(z)$ is real and nonnegative which is impossible for a general complex-valued function $f(z)$. However, this is possible when $f(z) = L'$ or $f(z)$ is a real-valued function with constant sign.

(iv) The signum function sgn on $\mathbb{C}$ is defined by

$$sgn\,(z) := \begin{cases} \dfrac{|z|}{z} & \text{for } z \neq 0 \\ 0 & \text{for } z = 0 \end{cases} = \begin{cases} \dfrac{\overline{z}}{|z|} & \text{for } z \neq 0 \\ 0 & \text{for } z = 0. \end{cases}$$

This function is clearly continuous on $\mathbb{C}\setminus\{0\}$ and

$$|sgn\,(z)| = \begin{cases} 1 & \text{for } z \neq 0 \\ 0 & \text{for } z = 0. \end{cases}$$ ●

A point z_0 in a set $D \subseteq \mathbb{C}$ that is not a limit point of D is called an *isolated point* of D. Clearly, at an isolated point z_0, there exists a $\delta > 0$ such that $N(z_0;\delta) \cap D = \{z_0\}$. A function $f : D \to \mathbb{C}$ is obviously continuous at all isolated points of D. For example, consider

$$f(z) = \begin{cases} z & \text{for } z \in \{1 - 1/n : \, n = 1, 2, \ldots\} \\ 1 & \text{for } z = 1 \end{cases}$$

and let $D = \{1 - \frac{1}{n} : \, n = 1, 2, \ldots\} \cup \{1\}$. The only limit point of D is 1 and so all other points of D are isolated. Since

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

f is continuous at $z = 1$. By definition, f is obviously continuous at the isolated points $z = 1 - 1/n$, $n = 1, 2, \ldots$. Thus, f is continuous on D.

What is the relationship between limits of sequences and limits of more general functions? A complex sequences $\{z_n\}_{n\geq 1}$, which defines a mapping $f : \mathbb{N} \to \mathbb{C}$, converges to z_0 if for every $\epsilon > 0$, there exists an $M > 0$ such that

$$f(N(\infty; M) \cap \mathbb{N}) \subset N(z_0; \epsilon).$$

Recall that a real M neighborhood of ∞ is the set of points outside the interval $(-M, M)$.

If the preimage of f is an unbounded region instead of the set of positive integers, we have the following analog: Let $f : \mathbb{C} \to \mathbb{C}$. Then $\lim_{z\to\infty} f(z) = L$ if for $\epsilon > 0$, there exists an $M > 0$ with $f(N(\infty; M)) \subset N(L; \epsilon)$.

Even if our region is bounded, there are important similarities between limits of the sequences and limits of more general functions. A sequence has a limit if eventually its points are "close" to one another, while a function of a complex variable has a limit if closeness of points in different planes is preserved. Our next theorem shows that continuity may be viewed as an operation that preserves convergence of sequences.

Theorem 2.36. *The function $f(z)$, defined in a region R, is continuous at a point $z_0 \in R$ if and only if, for every sequence $\{z_n\}$ in R converging to z_0, the sequence $\{f(z_n)\}$ converges to $f(z_0)$.*

Proof. Let $f(z)$ be continuous at z_0. Then, for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(z) - f(z_0)| < \epsilon$ whenever $|z - z_0| < \delta$ $(z \in R)$. If $\{z_n\}$ converges to z_0, then $|z_n - z_0| < \delta$ for $n > N$. By continuity, $|f(z_n) - f(z_0)| < \epsilon$ for $n > N$. Since ϵ was arbitrary, the sequence $\{f(z_n)\}$ converges to $f(z_0)$.

Conversely, suppose that $f(z)$ is not continuous at z_0. Now discontinuity of f at z_0 means that (see (2.9)) for some $\epsilon > 0$, $N(f(z_0); \epsilon)$ does not contain the image of any neighborhood of z_0. This means that we can find a sequence of points $\{z_n\}$ such that $z_n \in N(z_0; 1/n) \cap R$ and $f(z_n) \notin N(f(z_0); \epsilon)$. As $|z_n - z_0| < 1/n$ for all n, the sequence $\{z_n\}$ converges to z_0 although the sequence $\{f(z_n)\}$ does not converge to $f(z_0)$. This contradiction completes the proof. ∎

Remark 2.37. Theorem 2.36 is equally valid for real-valued functions of a real variable. ●

Let f be a continuous function defined in a region A. What properties of A are inherited by its image $f(A)$? Theorem 2.36 states that convergent sequences in A give rise to convergent sequences in $f(A)$. But many properties, even for real-valued functions of a real variable, are not preserved under a continuous map.

Examples 2.38. (i) The function $f(z) = |z|$ maps the plane onto the real interval $[0, \infty)$. This shows that the continuous image of an open set need not be open. We then say that f is not an open map.

(ii) The function $f(x) = \tan^{-1} x$ maps the real line onto $(-\pi/2, \pi/2)$. This shows that the continuous image of a closed set need not be closed.

(iii) The function $f(z) = 1/z$ maps the punctured disk $0 < |z| < 1$ onto the exterior of the unit disk. This shows that the continuous image of a bounded set need not be bounded. ●

But all is not lost. If we combine the "nice" properties of the last two examples, the image must also be "nice".

Theorem 2.39. *The continuous image of a compact set is compact.*

Proof. Let $f : A \to f(A)$ be continuous on the compact set A. For any sequence $\{w_n\}$ in $f(A)$, we can find a corresponding sequence $\{z_n\}$ in A such that $f(z_n) = w_n$. By Theorem 2.27, there exists a subsequence $\{z_{n_k}\}$ that converges to a point $z_0 \in A$. By Theorem 2.36, $f(z_{n_k}) = w_{n_k}$ converges to a point $f(z_0) \in f(A)$. Since $\{w_n\}$ was arbitrary, every sequence in $f(A)$ has a subsequence that converges in $f(A)$. Hence $f(A)$ must be a compact set. ∎

A function f is said to be *locally constant* if for each $a \in D$ there exists a neighborhood $N(a; \delta)$ of a on which $f(z) = f(a)$ for all z.

Theorem 2.40. *If a continuous function on a connected set D is locally constant, then f is constant throughout.*

Proof. Let a be such that $f(a) = b$. Define

$$S = \{z : f(z) = b\} = f^{-1}(b).$$

Now S is open because f is locally continuous. But S is closed because the singleton set $\{b\}$ is closed. Since S is not empty, we must have $D = S$. This completes the proof. ∎

Because the complex field is not ordered, it makes no sense to talk about maximum and minimum values for a complex-valued function $f(z)$. However, the next best thing is a discussion of maxima and minima for the related real-valued function $|f(z)|$. It will be helpful to observe that $|f(z)|$ is continuous in any region where $f(z)$ is continuous. This follows from the inequality

$$\big|\,|f(z_2)| - |f(z_1)|\,\big| \le |f(z_2) - f(z_1)| \quad (z_1, z_2 \in \mathbb{C}).$$

Theorem 2.41. *If $f(z)$ is continuous on a compact set E, then $|f(z)|$ attains a maximum and minimum on E.*

Proof. According to Theorem 2.39, the image of E under $|f(z)|$, which we shall denote by E', is a compact set. Since E' is a bounded set of real numbers, it has a least upper bound b. As a consequence of Exercise 2.29(2), the point b is in the set E'. But this means that $|f(z_0)| = b$ for some $z_0 \in E$.

The proof that $|f(z)|$ attains its minimum is similar, with greatest lower bound being substituted for least upper bound. ∎

A function $f(z)$ is said to *uniformly continuous* in a region R if for every $\epsilon > 0$, there exists a $\delta > 0$ (δ depending only on ϵ) such that if $z_1, z_2 \in R$ and $|z_1 - z_2| < \delta$, then $|f(z_1) - f(z_2)| < \epsilon$. This differs from continuity in a region in that the same δ may be used for every point in the region.

For example, the function $f(z) = z$ is uniformly continuous in every region, since we may always choose $\delta = \epsilon$.

Examples 2.42. The function $f(z) = 1/z$, although continuous, is not uniformly continuous in the region $0 < |z| < 1$.

To see this, assume that $f(z)$ is uniformly continuous. Then for $\epsilon > 0$ there exists a δ, $0 < \delta < 1$, to satisfy the conditions of the definition. We exploit the sensitivity of this function near the origin. Let $z_1 = \delta$ and $z_2 = \delta/(1+\epsilon)$. Then $|z_1 - z_2| = \delta\epsilon/(1+\epsilon) < \delta$, but

$$|f(z_1) - f(z_2)| = \left|\frac{1}{\delta} - \frac{1+\epsilon}{\delta}\right| = \frac{\epsilon}{\delta} > \epsilon,$$

showing that f is not uniformly continuous on the punctured unit disk.

Here is another example. The function $f(z) = z^2$ is not uniformly continuous in the complex plane $\mathbb{C}$.

Again, assume the contrary and let $\epsilon > 0$ be given. Then for any $\delta > 0$, choose

$$z_1 = 1/\delta \quad \text{and} \quad z_2 = 1/\delta + \delta/(1+\epsilon).$$

Then, we have $|z_1 - z_2| = \delta/(1+\epsilon) < \delta$ and

$$|z_1^2 - z_2^2| = 2/(1+\epsilon) + \delta^2/(1+\epsilon)^2 > 2/(1+\epsilon).$$

Note that this function is uniformly continuous in any bounded region. ●

Example 2.43. Consider $f(z) = x^2 - iy^2$. Clearly f is continuous on $\mathbb{C}$. But f is not uniformly continuous on $\mathbb{C}$, whereas it is uniformly continuous for $|z| < R$. To verify the second part we first note that, for $z = x + iy$ and $z_0 = x_0 + iy_0$,

$$\begin{aligned} |f(z) - f(z_0)| &= |(x + x_0)(x - x_0) - i(y + y_0)(y - y_0)| \\ &\leq |x + x_0|\,|x - x_0| + |y + y_0|\,|y - y_0|. \end{aligned}$$

If z, z_0 are in the disk $|z| < R$, then $|x + x_0| < 2R$ and $|y + y_0| < 2R$. This implies that

$$|f(z) - f(z_0)| \leq 2R[|x - x_0| + |y - y_0|] \leq 2\sqrt{2}R|z - z_0|$$

(since $|x| + |y| \leq \sqrt{2}|z|$). Now, given any $\epsilon > 0$, there exists a $\delta = \epsilon/(2\sqrt{2}R)$ such that

$$|f(z) - f(z_0)| < \epsilon \quad \text{whenever} \quad |z - z_0| < \delta = \frac{\epsilon}{2\sqrt{2}R}.$$

So, f is uniformly continuous on Δ_R.

The first part may now be verified as in the previous two examples, and so we leave this part as a simple exercise. ●

Theorem 2.44. *If $f(z)$ is continuous on a compact set A, then $f(z)$ is uniformly continuous on A.*

Proof. Let $\epsilon > 0$ be given. Then, for each point $z_\alpha \in A$, there is a neighborhood (depending on ϵ and z_α) such that

$$|f(z) - f(z_\alpha)| < \frac{\epsilon}{2} \tag{2.10}$$

whenever $|z - z_\alpha| < \delta_\alpha$, $z \in A$. The collection of all neighborhoods of the form $N(z_\alpha; \delta_\alpha/2)$ is a cover of A. By the compactness of A, there exists a finite subcover, say

$$A \subset \bigcup_{k=1}^{n} N\left(z_k; \frac{\delta_k}{2}\right). \tag{2.11}$$

Choose

$$\delta = \min\left\{\frac{\delta_1}{2}, \frac{\delta_2}{2}, \ldots, \frac{\delta_n}{2}\right\}.$$

We wish to show that this δ will work for the whole set A.

Let w_1 and w_2 be any two points in A such that $|w_1 - w_2| < \delta$. By (2.11), $w_1 \in N(z_k; \delta_k/2)$ for some k. According to (2.10), it follows that

$$|f(w_1) - f(z_k)| < \frac{\epsilon}{2}. \tag{2.12}$$

But we also have

$$|w_2 - z_k| \le |w_2 - w_1| + |w_1 - z_k| < \delta + \frac{\delta_k}{2} < \frac{\delta_k}{2} + \frac{\delta_k}{2} = \delta_k.$$

Hence $w_2 \in N(z_k; \delta_k) \cap A$ and

$$|f(w_2) - f(z_k)| < \frac{\epsilon}{2}. \tag{2.13}$$

Combining (2.12)and (2.13) we obtain

$$|f(w_1) - f(w_2)| \le |f(w_1) - f(z_k)| + |f(z_k) - f(w_2)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

and this completes the proof. ∎

We end the section with a remark on stereographic projection discussed in the previous section. If $\pi : S \setminus \{(0,0,1)\} \to \mathbb{C}$ is a function, then, according to the rule of correspondence,

$$\pi(x, y, u) = \left(\frac{x}{1-u}, \frac{y}{1-u}, 0\right) = \left(\frac{x}{1-u}\right) + i\left(\frac{y}{1-u}\right)$$

and π has an inverse function $\pi^{-1} : \mathbb{C} \to S \setminus \{(0,0,1)\}$ with the rule of correspondence

$$\pi^{-1}(z) = \left(\frac{2x}{|z|^2+1}, \frac{2y}{|z|^2+1}, \frac{|z|^2-1}{|z|^2+1} \right).$$

Thus, we have established the one-to-one correspondence between the Riemann sphere minus the north pole, namely $S \setminus \{(0,0,1)\}$, and $\mathbb{C}$. From these two formulas, it is evident that π and π^{-1} are continuous functions. In other words, the mapping π defined above is a homeomorphism, i.e., π is one-to-one onto, with both π and π^{-1} continuous. By allowing $(0,0,1)$ to map onto the point at infinity, it is evident that π maps S one-to-one onto $\mathbb{C}_\infty$. Moreover, if $s_1 = (x_1, y_1, u_1)$ and $s_2 = (x_2, y_2, u_2)$ are two points in S, then we define the distance function $d : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ by the Euclidean distance

$$\begin{aligned} d(s_1, s_2) &= |(x_1, y_1, u_1) - (x_2, y_2, u_2)| \\ &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (u_1 - u_2)^2}\,. \end{aligned}$$

Suppose now that s_1, s_2 are the images under the stereographic projection of $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ and define $\chi : \mathbb{C}_\infty \times \mathbb{C}_\infty \to \mathbb{R}$ by

$$\chi(z_1, z_2) = d(s_1, s_2).$$

Then it is easy to verify that χ is a metric on $\mathbb{C} \cup \{\infty\}$. We call χ the *chordal metric* on $\mathbb{C} \cup \{\infty\}$ and $(\mathbb{C}_\infty, \chi)$ the extended complex plane which is indeed isometric (i.e., distance preserveness) with (S, d). A straightforward exercise shows that the chordal distance is

$$\chi(z_1, z_2) = \begin{cases} \dfrac{2|z_1 - z_2|}{\sqrt{1+|z_1|^2}\sqrt{1+|z_2|^2}} & \text{if } z_1, z_2 \in \mathbb{C} \\ \dfrac{2}{\sqrt{1+|z_1|^2}} & \text{if } z_1 \in \mathbb{C}, z_2 = \infty \\ 0 & \text{if } z_1 = \infty,\ z_2 = \infty. \end{cases}$$

Let us see what the open disks look like that are centered at the point at infinity. A deleted ϵ-neighborhood of ∞ in $(\mathbb{C}_\infty, \chi)$ has the form

$$N_\chi(\infty; \epsilon) = \{z : \chi(z, \infty) < \epsilon\}.$$

According to the above formula

$$\chi(z, \infty) < \epsilon \iff (1 + |z|^2)^{-1/2} < \epsilon/2 \iff 1 + |z|^2 > (2/\epsilon)^2.$$

Assuming $\epsilon < 2$, this means $|z| > \sqrt{(2/\epsilon)^2 - 1}$. This shows that a deleted neighborhood of ∞ in $\mathbb{C}_\infty$ is of the form

$$N'_\chi(\infty; R) = \{z \in \mathbb{C} : |z| > R\}, \quad R > 0.$$

Note that if $\epsilon \geq 2$, $N_\chi(\infty; \epsilon) = \mathbb{C}_\infty$.

Finally, we now briefly indicate certain concepts associated with $(\mathbb{C}_\infty, \chi)$. A sequence $\{z_n\}$ in $\mathbb{C}$ converges to ∞ in $(\mathbb{C}_\infty, \chi)$ if and only if given $R > 0$

there exists an index $N = N(R)$ such that $|z| > R$ for all $n \geq N$. Similarly, if $f : A \to \mathbb{C}$ and $z_0 \in \mathbb{C}$ is a limit point of A, then $\lim_{z \to z_0} f(z) = \infty$ iff given $R > 0$ there exists a $\delta > 0$ such that $|f(z)| > R$ whenever $0 < |z - z_0| < \delta$ and $z \in A$. If ∞ is a limit point of A, then $\lim_{z \to \infty} f(z) = L$ iff given $\epsilon > 0$ there exists an $R > 0$ such that $|f(z) - L| < \epsilon$ whenever $|z| > R$ and $z \in A$.

Questions 2.45.

1. What ambiguities might there be if we called the preimage the domain of the function?
2. What is the geometric significance of a complex-valued function of a real variable? A real-valued function of a complex variable?
3. What properties do functions and their inverses have in common?
4. For what kinds of functions will we have points closer (more distant) in the w plane than in the z plane?
5. What can we say about the continuity of sequences?
6. Can we talk about a function being continuous at ∞?
7. What can we say about the continuous image of a limit point of a set?
8. How do the proofs of Theorem 2.36 and Theorem 2.18 compare?
9. What is the largest region on which $f(z) = 1/z$ is uniformly continuous?
10. Can discontinuous functions map compact sets onto compact sets?
11. If a function is uniformly continuous on a set A, is it also uniformly continuous on every subset of A?
12. How can you define a piecewise continuous real-valued function of a real variable defined on an interval $[a, b]$?
13. How can you define a piecewise continuous complex function of a real variable defined on an interval $[a, b]$?
14. Are piecewise continuous real-valued functions of a real variable defined on an interval $[a, b]$ integrable and bounded?
15. Does $f(z) = \arg z$ define a complex function? How about

$$f(z) = \cos(\arg z) + i \sin(\arg z)?$$

Exercises 2.46.

1. Find the following limits when they exist:

(a) $\displaystyle\lim_{z \to 3i} \frac{z^2 + 9}{z - 3i}$ (b) $\displaystyle\lim_{z \to 2i} \frac{\overline{z} + z^2}{1 - \overline{z}}$

(c) $\displaystyle\lim_{z \to \infty} \frac{z + 1}{z^2}$ (d) $\displaystyle\lim_{z \to \infty} \frac{z^2 + 10z + 2}{2z^2 - 11z - 6}$

(e) $\displaystyle\lim_{z \to 3i} \frac{z^3 + 27i}{z^2 + 9}$ (f) $\displaystyle\lim_{z \to 1} \frac{1 - z^n}{z^2 + 5z - 6} \quad (n \geq 1)$

2. Discuss continuity and uniform continuity for the following functions.

(a) $f(z) = \dfrac{1}{1-z}$ $(|z| < 1)$ (b) $f(z) = \dfrac{1}{z}$ $(|z| \geq 1)$

(c) $f(z) = \begin{cases} \dfrac{|z|}{z} & \text{if } 0 < |z| \leq 1 \\ 0 & \text{if } z = 0 \end{cases}$ (d) $f(z) = \begin{cases} \dfrac{\operatorname{Re} z}{z} & \text{if } 0 < |z| < 1 \\ 1 & \text{if } z = 0. \end{cases}$

3. Prove that $f(z) = 1/(1-z)$ is not uniformly continuous for $|z| < 1$.
4. Show that the function $f(z) = 1/z^2$ is not uniformly continuous for $0 < \operatorname{Re} z < 1/2$ but is uniformly continuous for $1/2 < \operatorname{Re} z < 1$.
5. Let $f(z)$ be one of the following functions each being defined in the punctured plane $\mathbb{C} \setminus \{0\}$:

$$\frac{\operatorname{Re} z}{z}, \quad \frac{\operatorname{Im} z}{z}, \quad \frac{z}{|z|}, \quad \frac{z}{\overline{z}}, \quad \frac{|z|}{z}, \quad \frac{\overline{z}}{z}.$$

Is it possible to suitably define any one of the these functions at $z = 0$ so that the resulting function will become continuous at $z = 0$. Answer the same question for the functions

$$\frac{z\operatorname{Re} z}{|z|} \quad \text{and} \quad \frac{z\operatorname{Im} z}{|z|}.$$

How about for the functions

$$\frac{z}{\operatorname{Re} z} \quad \text{and} \quad \frac{z}{\operatorname{Im} z}$$

when it is defined for $\mathbb{C} \setminus \{x + iy : x \neq 0\}$ and $\mathbb{C} \setminus \{x + iy : y \neq 0\}$, respectively?
6. Discuss continuity of

$$f(z) = \begin{cases} \dfrac{(\operatorname{Re} z)^2\,(\operatorname{Im} z)}{|z|^2} & \text{if } z \neq 0 \\ 0 & \text{if } z = 0 \end{cases}$$

at the all points of $\mathbb{C}$.
7. Find the following limits:
(a) $\lim\limits_{z\to 0} f(z)$, where $f(z) = \dfrac{xy}{x^2+y^2} + 2xi$,
(b) $\lim\limits_{z\to 0} f(z)$, where $f(z) = \dfrac{xy}{x^2+y} + 2\dfrac{x}{y}i$,
(c) $\lim\limits_{z\to 0} f(z)$, where $f(z) = \dfrac{xy^3}{x^3+y^3} + \dfrac{x^8}{y^2+1}i$.
8. If $\lim_{z\to\infty} f(z) = a$, and $f(z)$ is defined for every positive integer n, prove that $\lim_{n\to\infty} f(n) = a$. Give an example to show that the converse is false.

9. Show that a monotonic real-valued function of a real variable cannot have uncountably many discontinuities.
10. Show that $f : A \to B$ is continuous if and only if for every open set O relative to B, $f^{-1}(O)$ is an open set relative to A.
11. Using Exercise 2.46(10), prove that the continuous image of a compact set is compact.
12. Show that $f : A \to B$ is continuous if and only if for every closed set F relative to B, $f^{-1}(F)$ is a closed set relative to A.
13. Prove that continuous image of a connected set is connected.
14. If a function, defined on a compact set, is continuous, one-to-one, and onto, show that the inverse function also has these properties. Can compactness be omitted?
15. Let f and g be continuous on a set A. Show that $f + g$, $f \cdot g$, and $f/g\,(g \neq 0)$ are also continuous on A. What can we say if f and g are uniformly continuous on A?
16. Show that $f(z)$ is continuous in a region R if and only if both $\operatorname{Re} f(z)$ and $\operatorname{Im} f(z)$ are continuous in R.
17. Show that every polynomial is continuous in the complex plane.
18. Let $f(z)$ be continuous in the complex plane. Let $A = \{z \in \mathbb{C} : f(z) = 0\}$. Show that A is a closed set.
19. Show that $\displaystyle\lim_{z\to 4} \frac{1}{z-4} = \infty$ and $\displaystyle\lim_{z\to\infty} \frac{1}{z^2+2} = 0$.
20. Suppose that $J : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is defined by $J(z) = 1/z$, $z \in \mathbb{C}_\infty$. Do our conventions imply $J(0) = \infty$ and $J(\infty) = \infty$? Does

$$\chi(J(z), J(w)) = \chi(z, w)$$

hold in $\mathbb{C}_\infty$?

3

Bilinear Transformations and Mappings

In the previous chapter we saw that a complex function of a complex variable maps points in the z plane onto points in the w plane. After the initial excitement of this discovery wore off, it became rather tiresome to map points onto points in computer like fashion. In this chapter we will see, for some special functions, what happens to regions in the z plane when mapped onto the regions in the w plane. We will show that bilinear transformations map circles and straight lines onto circles and straight lines. In fact, we will discover that—contrary to popular belief—a circle is very similar to a straight line, at least in the extended complex plane. We also determine the most general form of bilinear transformation which maps

- the real line $\mathbb{R}$ onto the unit circle $|z| = 1$
- the unit circle $|z| = 1$ onto itself
- the unit circle $|z| = 1$ onto $\mathbb{R}$
- the real line $\mathbb{R}$ onto itself.

3.1 Basic Mappings

The function $w = f(z) = z + b$, where b is a complex constant, maps sets in the z plane onto sets in the w plane displaced through a vector b. This mapping is known as a *translation.* Note that the set in the w plane will have the same shape and size as the set in the z plane. For instance, the function $w = z + (1 + 2i)$ maps the square having vertices $\pm 1 \pm i$ onto a square having vertices $i, 2 + i, 3i$, and $2 + 3i$ (see Figure 3.1). To show this, let $z = x + iy$ and $w = u + iv$. Then

$$u + iv = (x + iy) + (1 + 2i), \quad \text{i.e. }, u = x + 1, \;\; v = y + 2.$$

As x describes the interval $[-1, 1]$, u describes the interval $[0, 2]$; as y describes the interval $[-1, 1]$, v describes the interval $[1, 3]$.

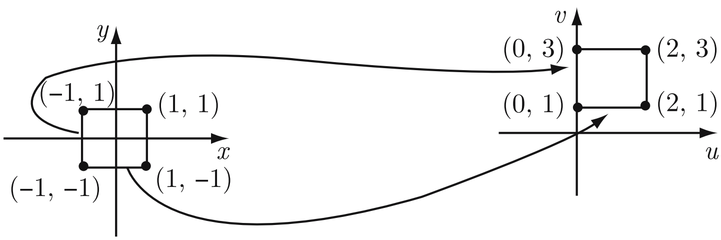


Figure 3.1. Image of a square under $w = z + 1 + 2i$

The function $w = az$, where $a = \cos\alpha + i\sin\alpha$, maps a point in the z plane onto a point in the w plane whose distance from the origin is the same but whose argument is increased by α, the argument of a. This mapping is called a *rotation.* For instance, the function $w = iz$ maps the right half-plane ($\operatorname{Re} z > 0$) onto the upper half-plane ($\operatorname{Im} z > 0$). Observing that $\operatorname{Arg} i = \pi/2$, we may view this geometrically as a mapping of the points in the z plane satisfying $-\pi/2 < \operatorname{Arg} z < \pi/2$ onto points in the w plane satisfying $0 < \operatorname{Arg} w < \pi$. Analytically,

$$w = u + iv = i(x + iy) = -y + ix, \quad \text{i.e.} \quad u = -y, \; v = x.$$

Thus $x > 0$ is mapped onto $v > 0$.

For $a > 0$, $a \neq 1$ the function $w = az$ is known as a *magnification* (although for $a < 1$ it is really a contraction). This function takes regions in the z plane and either stretches or shrinks them, depending on whether $a > 1$ or $a < 1$. For instance, the function $w = 5z$ maps the disk $|z| \leq 1$ onto the disk $|w| \leq 5$.

More generally, for complex values of a, the function $w = az$ represents both a rotation and a magnification; the expression $\arg a$ is the rotation part, and $|a|$ is the magnification part. Indeed, we can combine a translation, rotation, and magnification to obtain the *linear function*

$$w = f(z) = az + b,$$

where a and b are complex constants. Note that

$$|w_1 - w_2| = |f(z_1) - f(z_2)| = |a|\,|z_1 - z_2|$$

so that the distance between any two points is multiplied by $|a|$. For instance, the function

$$w = (1 - i)z + (2 + i)$$

maps the rectangle in the z plane shown in Figure 3.2 onto the rectangle in the w plane that has twice the area, with the length of each side being increased by a factor of $\sqrt{2}$.

There is a relationship between a complex linear function and the more familiar real-valued linear function $y = ax + b$, a straight line. The complex-valued function $w = az + b$, with a and b are complex constants, maps straight

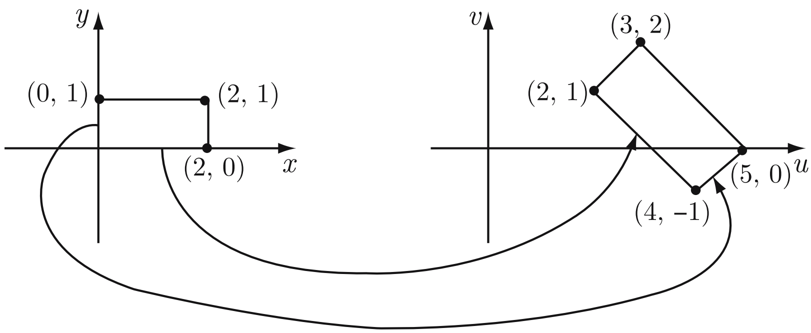


Figure 3.2. Image of a rectangle under $w = (1 - i)z + (2 + i)$

lines in the z plane onto straight lines in the w plane. Note that the complex linear functions ($a \neq 0$) always map ∞ to ∞. We leave the determination of the effect of the constants a and b on the slope of the image line as an exercise for the reader. Observe that $w = az + b$, like its real-valued counterpart, is a one-to-one function.

The mapping $w = 1/z$, called an *inversion*, takes points close to the origin in the z plane onto points far from the origin in the w plane and points far from the origin in the z plane onto points close to the origin in the w plane. Indeed if $z = re^{i\theta}$, then

$$w = \frac{1}{z} = \frac{1}{r}e^{-i\theta}.$$

In particular, as z approaches the origin, w approaches the point at ∞ in the extended complex plane; i.e., given $M > 0$, there exists a $\delta > 0$ such that $|z| < \delta$ implies $|w| > M$. We thus have a one-to-one map from the extended plane onto itself with the origin being mapped onto the point at ∞. However, it is wrong to conclude that inversion always maps lines into lines, and circles into circles (see Theorem 3.1)

There is also a certain symmetry with respect to both the unit circle and the real axis. Points inside (outside) the unit circle are mapped onto points outside (inside) the unit circle, and points above (below) the real axis are mapped onto points below (above) the real axis (see Figure 3.3).

The inversion $w = 1/z$ is sometimes called a *reflection* with respect to both the unit circle and the real axis. To see what happens to sets in the z plane when transformed into sets in the w plane by this reflection, we solve

$$w = u + iv = \frac{1}{z} = \frac{1}{x + iy}$$

for a given variable in one plane in terms of the variables in the other plane. This gives the relations

$$u = \frac{x}{x^2 + y^2}, \quad v = -\frac{y}{x^2 + y^2} \tag{3.1}$$

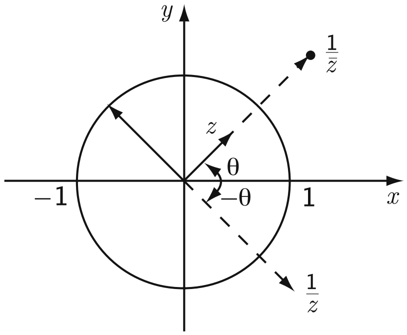


Figure 3.3. Illustration for $w = 1/z$

and

$$x = \frac{u}{u^2 + v^2}, \quad y = -\frac{v}{u^2 + v^2}. \tag{3.2}$$

From (3.1) or (3.2) we obtain

$$x^2 + y^2 = \frac{1}{u^2 + v^2}. \tag{3.3}$$

Now consider the equation

$$a(x^2 + y^2) + bx + cy + d = 0, \tag{3.4}$$

where a, b, c, and d are real constants. This equation represents a circle if $a \neq 0$ and a straight line if $a = 0$. From (3.1), (3.2), and (3.3), we see that the function $w = 1/z$ maps (3.4) onto the set

$$d(u^2 + v^2) + bu - cv + a = 0, \tag{3.5}$$

which describes a circle for $d \neq 0$ and a straight line if $d = 0$.

We can now, in view of (3.4) and (3.5), draw several conclusions about the mapping properties of $w = 1/z$:

(a) Circles not passing through the origin (that is, with $a \neq 0$ and $d \neq 0$). are mapped onto circles not passing through the origin.
(b) Circles passing through the origin (that is, with $a \neq 0$ and $d = 0$) are mapped onto straight lines not passing through the origin.
(c) Straight lines not passing through the origin (that is, with $a = 0$ and $d \neq 0$) are mapped onto circles passing through the origin.
(d) Straight line passing through the origin (that is, with $a = 0$ and $d = 0$) are mapped onto straight lines passing through the origin.
(e) The circle $|z| = 1$ maps onto the circle $|w| = 1$.
(f) The punctured disk $\varDelta \setminus \{0\}$ maps onto $\mathbb{C} \setminus \overline{\varDelta}$, and conversely.
(g) All points on $\mathbb{C} \setminus \overline{\varDelta}$ map onto $\varDelta \setminus \{0\}$.

In short, we have

Theorem 3.1. *The function $w = 1/z$ maps circles and straight lines onto circles and straight lines.*

Is there a way to remember which maps onto which? The key lies in the fact that the origin maps onto the point at ∞. Every straight line (and no circle) passes through the point at ∞. Hence a straight line or circle maps onto a straight line if it passes through the origin, and onto a circle if it does not. Also, we note that the interior of a circle containing the origin maps onto the exterior of a circle, and the interior of a circle not containing the origin (nor having the origin as a boundary point) maps onto the interior of a circle. Finally, we present some precise mapping properties of $w = 1/z$. Consider the circle $|z-a| = R$, $a \neq 0$. If $w = 1/z$, then we obtain that

$$
\begin{aligned}
|z-a| < R \iff \left|\frac{1}{w} - a\right| < R &\iff |1-aw|^2 < R^2|w|^2 \\
&\iff |w|^2(|a|^2 - R^2) - 2\mathrm{Re}\,(aw) + 1 < 0. \\
&\iff \begin{cases} \mathrm{Re}\,(aw) > 1/2 & \text{for } R = |a| \\ \left|w - \dfrac{\overline{a}}{|a|^2 - R^2}\right| < \dfrac{R}{|a|^2 - R^2} & \text{for } R < |a| \\ \left|w - \dfrac{\overline{a}}{|a|^2 - R^2}\right| > \dfrac{R}{R^2 - |a|^2} & \text{for } R > |a|. \end{cases}
\end{aligned}
$$

For example, if $R = |a|$, then under the inversion $w = 1/z$ we have

- $|z-a| < |a|$ is mapped onto the half-plane $\mathrm{Re}\,(aw) > 1/2$
- $|z-a| = |a|$ is mapped onto the straight line $\mathrm{Re}\,(aw) = 1/2$
- $|z-a| > |a|$ is mapped onto the half-plane $\mathrm{Re}\,(aw) < 1/2$.

When $|a|^2 - R^2 \neq 0$, there exist two possibilities $|a| > R$ and $|a| < R$. In each of these cases, mapping properties may be stated with the help of the above discussion. For example, under the inversion $w = 1/z$, we have the following:

- $|z-3| < R$ is mapped onto the disk $\left|w - \dfrac{3}{9-R^2}\right| < \dfrac{R}{9-R^2}$ for $R < 3$
- $|z-3| < R$ is mapped onto the disk $\left|w + \dfrac{3}{R^2-9}\right| > \dfrac{R}{R^2-9}$ for $R > 3$
- $|z-3| = R$ is mapped onto the disk $\left|w - \dfrac{3}{9-R^2}\right| = \dfrac{R}{|9-R^2|}$ for $R \neq 3$.

Questions 3.2.

1. The functions $w = 1/z$, $w = \overline{z}$, and $w = -z$ all map the upper half of the unit circle onto lower half. What are their differences?

2. How does the area of a region compare with the area of its image for a linear function? For an inversion?
3. For $w = 1/z$, what is the image of the interior of a circle having the origin as boundary point?
4. What happens to conic sections other than circles under an inversion?
5. Is there a difference between performing a translation followed by an inversion and an inversion followed by a translation?
6. For the four operations of translation, rotation, magnification, and inversion, which pairs may be interchanged without affecting the mapping properties?

Exercises 3.3.

1. For the mapping $w = (1+i)z + 2$, find the image of
   (a) the line $y = 2x$ (b) the line $y = 3x + 2$
   (c) the circle $|z| = 3$ (d) the circle $|z-1| = 2$.
2. Find the image of the half-plane $\mathrm{Re}\, z > 0$ under the transformation
   (a) $w = 2iz - i$ (b) $w = i/z - 1$.
3. Find the image of the semi-infinite strip $\{z : 0 < \mathrm{Re}\, z < 2,\ \mathrm{Im}\, z > 1\}$ for the transformation $w = (1-i)z + (2-i)$, and sketch.
4. Find a linear transformation f that maps the circle $|z+1| = 2$ onto the circle $|w+i| = 3$. Find also the image of $|z+1| < 2$ under f.
5. Prove that the linear transformation $w = az + b$ maps a circle having radius r and center z_0 onto a circle having radius $|a|r$ and center $az_0 + b$.
6. Given a triangle with vertices at $3+4i$, $-3+4i$, and $-5i$, find its image for the transformation
   (a) $w = z + 5i$ (b) $w = iz + (2-i)$ (c) $w = (2+i)z - 3$.
7. Find the image of the line $y = 2x + 1$ under the following transformations.
   (a) $w = 1/z$ (b) $w = i/z$ (c) $w = 1/(z-2i)$.
8. For the transformation $w = 1/z$, find the image of
   (a) the circle $|z-2| = 1$ (b) the circle $|z-1| = 2$
   (c) the circle $|z-1| = 1$ (d) the domain $\mathrm{Re}\, z > 1$.
   (e) the infinite strip $\frac{1}{4} < \mathrm{Re}\, z < \frac{1}{2}$.
9. Find the images of the strips $0 \le \mathrm{Re}\, z \le 2$ and $0 \le \mathrm{Im}\, z \le 2$ under the map $w = 1/z$.

3.2 Linear Fractional Transformations

By considering quotient of two linear transformations, we get a very important class of mappings of the form, known as a *linear fractional transformation*:

$$w = T(z) = \frac{az+b}{cz+d}, \tag{3.6}$$

where a, b, c and d are the complex numbers such that $ad - bc \neq 0$. This transformation contains, as special cases, the mappings of the previous section. For $c = 0$ we have a linear transformation, and for $a = d = 0$, $b = c$ we have an inversion. The condition $ad - bc \neq 0$ ensures that the mapping is not a constant. To see this, suppose that $ad - bc = 0$. If $c \neq 0$, we may solve for $b = ad/c$ and write (3.6) as

$$w = \frac{\frac{a}{c}(cz+d) + b - \frac{ad}{c}}{cz+d} = \frac{a}{c} + \frac{b - \frac{ad}{c}}{cz+d} = \frac{a}{c}, \tag{3.7}$$

a constant. Similarly if $a \neq 0$, then $ad - bc = 0$ gives $d = bc/a$ so that (3.6) becomes

$$w = \frac{az+b}{cz+bc/a} = \frac{a}{c}\left(\frac{z+b/a}{z+b/a}\right) = \frac{a}{c},$$

again a constant. If $ad - bc = 0$ and $a = 0$, then either $b = 0$ or $c = 0$. When $a = b = 0$, it follows that $w = 0$, and when $a = c = 0$, we see that $w = b/d$, a constant. We will henceforth assume that $ad - bc \neq 0$. Thus, we write

$$w = T(z) = \frac{az+b}{cz+d} = \begin{cases} \frac{a}{c} - \left(\frac{ad-bc}{c^2}\right)\frac{1}{z+d/c} & \text{if } c \neq 0 \\ \left(\frac{a}{d}\right) z + \frac{b}{d} & \text{if } c = 0. \end{cases} \tag{3.8}$$

The domain of the definition of $T(z)$ is $\mathbb{C} \setminus \{-d/c\}$. Clearly, $T(z)$ is a one-to-one function on its domain. Since T is well defined for all points in the extended complex plane except at $z = -d/c$ and the point at ∞, we may extend the definition of T to the extended complex plane by including these points. Indeed, as

$$\lim_{z \to -d/c} \frac{1}{T(z)} = \lim_{z \to -d/c} \frac{cz+d}{az+b} = \frac{0}{a(\frac{-d}{c})+b} = 0,$$

we find that $\lim_{z \to -d/c} T(z) = \infty$. Further, we have

$$\lim_{z \to \infty} \frac{az+b}{cz+d} = \lim_{z \to 0} \frac{\frac{a}{z}+b}{\frac{c}{z}+d} = \lim_{z \to 0} \frac{a+bz}{c+dz} = \frac{a}{c},$$

and hence for $c \neq 0$, we may define

$$T(z) = \begin{cases} \frac{az+b}{cz+d} & \text{if } z \neq -d/c,\ z \neq \infty \\ \infty & \text{if } z = -d/c \\ \frac{a}{d} & \text{if } z = \infty \end{cases}$$

and T defined in this way is then one-to-one onto the extended complex plane and has an inverse that is also a linear fractional transformation. Solving for z, in terms of w, we obtain

$$z = T^{-1}(w) = \begin{cases} \dfrac{dw-b}{-cw+a} & \text{if } w \neq a/c,\ w \neq \infty \\ \infty & \text{if } w = a/c \\ -\dfrac{d}{c} & \text{if } w = \infty. \end{cases}$$

Thus regions may be mapped with equal facility from the extended z plane to the extended w plane or from the extended w plane to the extended z plane. When cleared of fractions, (3.6) assumes the form

$$Azw + Bz + Cw + D = 0,$$

an equation that is linear in both z and w. For this reason, a linear fractional transformation is often called a *bilinear transformation.*

A bilinear transformation represents a one-to-one continuous mapping of the extended complex plane onto itself with the point $z = -d/c$ mapping onto $w = \infty$ and the point $z = \infty$ mapping onto $w = a/c$. Recall that

- a linear transformation maps circles onto circles and straight lines onto straight lines.
- an inversion maps circles and straight lines onto circles and straight lines, see Theorem 3.1.

We will use these facts to deduce that bilinear transformations have similar mapping properties. We consider $w = T(z)$ defined by (3.8) for $c \neq 0$. Then we have that

$$w = T(z) = (f_3 \circ f_2 \circ f_1)(z),$$

where

$$w_1 = f_1(z) = z + \frac{d}{c}, \quad w_2 = f_2(w_1) = \frac{1}{w_1}$$

and

$$w = f_3(w_2) = \frac{a}{c} - \left(\frac{ad-bc}{c^2}\right) w_2.$$

That is $z \mapsto w = T(z)$ is given by the composition

$$z \mapsto w_1 \mapsto w_2 \mapsto w.$$

The first is linear and maps circles in the z plane onto circles in the w_1 plane and straight line in the z plane onto straight lines in the w_1 plane; the second is an inversion, mapping circles and straight lines in the w_1 plane onto circles and straight lines in the w_2 plane; the third is again linear and maps circles in the w_2 plane onto circles in the w plane and straight lines in the w_2 plane onto straight lines in the w plane. If $c = 0$, the transformation is a linear. The above results may be summarized as

Theorem 3.4. *The bilinear transformation maps circles and straight lines onto circles and straight lines.*

In the foregoing proof, we have also shown that a linear fractional transformation (3.8) can be written as a composition of three types of elementary transformations namely,

- the translation $T_1(z) = z + B$
- the inversion $T_2(z) = 1/z$
- the dilation $T_3(z) = Az$.

Note that a dilation is a composition of a magnification (or contraction) and rotation. The order in which these transformations are performed is immaterial as they commute.

Remark 3.5. The point $z = -d/c$ plays the same role in the bilinear transformation as does the point $z = 0$ in the inversion transformation. Thus a straight line or a circle maps onto a straight line if it passes through the point $z = -d/c$, and onto a circle if it does not. •

Theorem 3.4 may often be used to simplify computation. For example, in Exercise 1.8(6), the reader was asked to show that

$$\mathrm{Re}\left\{\frac{z}{1-z}\right\} > -\frac{1}{2} \quad \text{for} \quad |z| < 1.$$

Presumably, the reader separated $z/(1-z)$ into its real and imaginary parts, substituted in points inside the unit disk, and then marveled at the result. We will now apply more sophisticated techniques that give some insight into the solution. The bilinear transformation $w = z/(1-z)$ maps the unit circle onto a straight line (since $z = 1$ maps onto $w = \infty$). By choosing any two distinct points on the unit circle, we can determine this straight line. The point -1 and i map onto the points $-\frac{1}{2}$ and $-\frac{1}{2} + \frac{i}{2}$, respectively. Thus the image of the circle $|z| = 1$ is the line $\mathrm{Re}\, w = -\frac{1}{2}$ (see Figure 3.4). The continuity of a bilinear map reveals that connected sets are mapped onto connected sets (see Exercise 2.46(13)). Since a bilinear map is also a one-to-one mapping of the extended plane onto itself, the image of $|z| < 1$ is either

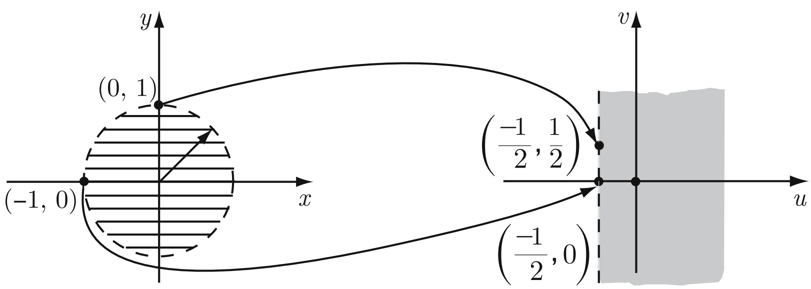


Figure 3.4. Illustration for the image of the unit disk under $z/(1-z)$

$$\mathrm{Re}\, w > -\frac{1}{2} \quad \text{or} \quad \mathrm{Re}\, w < -\frac{1}{2}.$$

To determine the correct image, we need to test only one point. The origin mapping onto itself assures us that $|z| < 1$ maps onto $\mathrm{Re}\, w > -\frac{1}{2}$. Alternatively, we simply note that

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

Thus, $|z| = 1$ gives

$$|w|^2 = |1+w|^2, \quad \text{i.e.,} \quad \mathrm{Re}\, w = -1/2$$

and, since 0 is mapped onto $w = 0$, it follows that $|z| < 1$ is mapped onto $\mathrm{Re}\, w > -1/2$ under the map $w = z/(1-z)$.

Example 3.6. Suppose the reader was asked to find the image of the closed half disk $\{z : |z| \leq 1, \mathrm{Re}\, z \geq 0\}$ under the bilinear transformation

$$w = \frac{z}{1-z}.$$

We know that $|z| < 1$ is mapped onto $\mathrm{Re}\, w > -1/2$. Moreover, $\mathrm{Re}\, z \geq 0$ is mapped onto points

$$\mathrm{Re}\, w + |w|^2 = |w + (1/2)|^2 - 1/4 \geq 0, \quad \text{i.e.,}\ |w + 1/2| \geq 1/2.$$

Consequently, the image of the closed half disk $\{z : |z| \leq 1, \mathrm{Re}\, z \geq 0\}$ under the bilinear transformation $w = z/(1-z)$ is

$$\{w : \mathrm{Re}\, w \geq -1/2\} \cap \{w : |w + 1/2| \geq 1/2\}.$$

Note also that $\mathrm{Im}\, z \leq 0$ is mapped onto $\mathrm{Im}\, w \leq 0$, showing that the image of the closed half disk $\{z : |z| \leq 1, \mathrm{Im}\, z \leq 0\}$ under $w = z/(1-z)$ is $\{w : \mathrm{Re}\, w \geq -1/2 \text{ and } \mathrm{Im}\, w \leq 0\}$. What is the image of one-quarter disk $\{z : |z| \leq 1, \mathrm{Im}\, z \leq 0, \mathrm{Re}\, z \geq 0\}$? ●

Example 3.7. Suppose the reader is asked to find the image of the annulus $\{z : 1 < |z| < 2\}$ under $w = z/(1-z)$. To do this, we first note that

$$z = \frac{w}{1+w}$$

and so

$$|z| > 1 \iff |w|^2 > |1+w|^2, \quad \text{i.e.,} \quad 0 > 1 + 2\mathrm{Re}\, w$$

and

$$|z| < 2 \iff |w|^2 < 4|1+w|^2, \quad \text{i.e.} \quad 0 < 3[|w + (4/3)|^2 - (4/9)].$$

So we see easily that

- $|z| > 1$ is mapped into $\operatorname{Re} w < -1/2$
- $|z| < 2$ is mapped onto $|w + 4/3| > 2/3$

and the required image is $\{w : \operatorname{Re} w < -1/2\} \cap \{w : |w + 4/3| > 2/3\}$. ●

So far, we have been concerned with determining the images of sets under a fixed bilinear transformation. We now reverse the problem and consider the following.

Problem 3.8. Given two sets, under what circumstances will there exist a bilinear map from one set onto the other?

From elementary geometry, we know that any three points determine either a circle or a straight line depending on whether the three points are collinear. Any bilinear transformation maps a circle or a straight line determined by these points onto either a circle or a straight line, depending on the coefficients of the bilinear transformation.

For our next result, we need the notion of fixed points. A point z in $\mathbb{C}_\infty$ that satisfies the equation

$$z = T(z)$$

is called a *fixed point* of T. The identity transformation $I(z) = z$ has every z in $\mathbb{C}_\infty$ as its fixed points. Concerning other bilinear transformations, we have the following result which has many important consequences.

Theorem 3.9. *A bilinear transformation $w = T(z)$ with more than two fixed points in $\mathbb{C}_\infty$ must be the identity transformation.*

Proof. Suppose that $c = 0$ in (3.8). Then, T is of the form

$$T(z) = \alpha z + \beta, \quad \alpha \neq 0.$$

The solution of $z = \alpha z + \beta$ are the fixed points of T. Clearly, the solution set is given by

(i) $z = \infty$ and $z = \beta/(1 - \alpha)$ whenever $\alpha \neq 1$
(ii) $z = \infty$ whenever $\alpha = 1, \beta \neq 0$
(iii) all z whenever $\alpha = 1, \beta = 0$.

Suppose that $c \neq 0$. Then ∞ cannot be a fixed point and the fixed point equation $z = T(z)$ gives the quadratic equation

$$cz^2 + (d - a)z - b = 0$$

which has at most two complex roots. Evidently, the only situation which provides more than two fixed points is the one in which $T = I$, the identity transformation. ■

We will now show that for A (a circle or a straight line in the z plane) and B (a circle or a straight line in the w plane), there exists a bilinear map from A onto B.

Theorem 3.10. *Given three distinct points, z_1, z_2, and z_3 in the extended z plane and three distinct points w_1, w_2, and w_3 in the extended w plane, there exists a unique bilinear transformation $w = T(z)$ such that $T(z_k) = w_k$ for $k = 1, 2, 3$.*

Proof. We first assume that none of the six points is ∞. Let T be given by (3.6). We wish to solve for a, b, c, and d in terms of z_1, z_2, z_3, w_1, w_2, and w_3. This sounds more complicated than it is. For $k = 1, 2, 3$, we have

$$w - w_k = \frac{az + b}{cz + d} - \frac{az_k + b}{cz_k + d} = \frac{(ad - bc)(z - z_k)}{(cz + d)(cz_k + d)}. \tag{3.9}$$

From (3.9) we obtain

$$\frac{w - w_1}{w - w_3} = \left(\frac{cz_3 + d}{cz_1 + d}\right)\left(\frac{z - z_1}{z - z_3}\right). \tag{3.10}$$

Replacing z by z_2 and w by w_2 in (3.10) leads to

$$\frac{w_2 - w_3}{w_2 - w_1} = \left(\frac{cz_1 + d}{cz_3 + d}\right)\left(\frac{z_2 - z_3}{z_2 - z_1}\right). \tag{3.11}$$

Multiplying (3.10) by (3.11) we have

$$\frac{(w - w_1)(w_2 - w_3)}{(w - w_3)(w_2 - w_1)} = \frac{(z - z_1)(z_2 - z_3)}{(z - z_3)(z_2 - z_1)}. \tag{3.12}$$

Solving for w in terms of z and the six points gives the desired transformation. If one of the points were the point at ∞, say $z_3 = \infty$, (3.12) would be modified by taking the limit as z_3 approached ∞. In this case, we would have

$$\frac{(w - w_1)(w_2 - w_3)}{(w - w_3)(w_2 - w_1)} = \frac{z - z_1}{z_2 - z_1}.$$

Now suppose that $S(z)$ and $T(z)$ are both bilinear transformations that agree at three or more points in $\mathbb{C}_\infty$, say

$$w_k = S(z_k) = T(z_k) \ \text{ for } k = 1, 2, 3.$$

Then for $k = 1, 2, 3$,

$$(S^{-1} \circ T)(z_k) = S^{-1}(T(z_k)) = S^{-1}(w_k) = z_k$$

and so, by Theorem 3.9, $S^{-1} \circ T = I$. This gives $S = T$ which proves the uniqueness part of the theorem. ∎

Corollary 3.11. *Given three distinct points, z_1, z_2, and z_3 in the extended z plane there exists a unique bilinear transformation $w = T(z)$ such that $T(z_1) = 0$, $T(z_2) = 1$, $T(z_3) = \infty$ and it is given by*

$$w = \frac{(z - z_1)(z_2 - z_3)}{(z - z_3)(z_2 - z_1)}.$$

Remark 3.12. The right side of (3.12) is called the *cross ratio* of the points z_1, z_2, z_3, z and is denoted by (z_1, z_2, z_3, z). Observe that (3.12) asserts the invariance of the cross ratio under a bilinear transformation. That is, for any four distinct pairs of points (z_1, w_1), (z_2, w_2), (z_3, w_3), and (z, w) of a bilinear transformation, we must have

$$(z_1, z_2, z_3, z) = (w_1, w_2, w_3, w).$$ •

Remark 3.13. It is not surprising that three points determine a bilinear transformation. If we divide (3.6) by one of the nonzero constants (assume $a \neq 0$), then (3.6) may be rewritten as

$$w = \frac{z + B}{Cz + D},$$

and elementary algebra may be used to solve for three equations with three unknowns. Moreover, the proof of Theorem 3.10 suggests a method of finding $w = T(z)$ satisfying the condition $w_j = T(z_j)$ whenever z_j and w_j are given, $j = 1, 2, 3$. In many cases, T can be found with even less trouble as we can seen in some of the examples of this section. •

Example 3.14. Let us now find a bilinear transformation that maps the points $z = i, 2, -2$ onto $w = i, 1, -1$, respectively.

To do this we may simply use (3.12) and obtain

$$\frac{(w - i)(1 + 1)}{(w + 1)(1 - i)} = \frac{(z - i)(2 + 2)}{(z + 2)(2 - i)}.$$

Solving this equation, we obtain

$$w = \frac{3z + 2i}{iz + 6}.$$

Similarly, it is easy to find the bilinear transformation that maps the points $z = 1 - i, 1 + i, -1 + i$ onto $0, 1, \infty$, respectively.

Indeed, by Corollary 3.11, we see that the desired transformation is

$$w = \frac{z - (1 - i)}{iz + (1 + i)}.$$ •

Example 3.15. Let us find a bilinear transformation which maps the disk $|z + i| < 1$ onto the exterior disk $|w| > 4$. To do this, we consider

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

Without loss of generality we may assume that $f(-i) = \infty$. Then $f(z)$ takes the form

$$f(z) = \frac{az + b}{z + i}.$$

Note that $f(0) = -ib$ and $f(-2i) = 2a + ib$. According to our requirements, these two points must lie on the circle $|w| = 4$. This gives

$$|b| = 4 \text{ and } |2a + ib| = \sqrt{4a^2 + b^2} = 4.$$

A choice satisfying these two conditions are $b = 4$ and $a = 0$. This shows that $f(z) = 4/(z + i)$ is a bilinear transformation which maps the circle $|z + i| = 1$ onto the circle $|w| = 4$. Since $-i \mapsto \infty$,

$$f(z) = \frac{4}{z + i}$$

is exactly a desired transformation. Note also that this is not unique as there are many bilinear transformations which do the same job. ●

Thus far we have seen that there is a unique bilinear transformation mapping three distinct points in the z plane onto three distinct points in the w plane. This has given rise to ways of mapping circles and straight lines onto circles and straight lines, although not uniquely. For example, to find a function mapping a line onto a circle, we may choose any three points on the line and make them correspond with any three points on the circle. Let $\operatorname{Im} z_0 > 0$. Then, by Theorem 3.10, the bilinear transformation mapping such that

$$z_0 \mapsto 0, \ \overline{z}_0 \mapsto \infty, \ 0 \mapsto \frac{z_0}{\overline{z}_0},$$

is given by

$$w = \frac{z - z_0}{z - \overline{z}_0}.$$

Note that this function maps the real line $\mathbb{R}$ onto the unit circle $|w| = 1$. Moreover, as $z_0 \mapsto 0$, it must map the upper half-plane $\{z : \operatorname{Im} z > 0\}$ onto the unit disk $|w| < 1$ and lower half-plane $\{z : \operatorname{Im} z < 0\}$ onto the exterior $|w| > 1$. On the other hand, the bilinear transformation mapping the points $z = \overline{z}_0, z_0, 0$ onto the points $w = 0, \infty, \overline{z}_0/z_0$ respectively, given by

$$w = \frac{z - \overline{z}_0}{z - z_0},$$

maps the lower half-plane onto the unit disk $|w| < 1$ and upper half-plane onto the exterior $|w| > 1$. How about the bilinear transformation such that

$$z_0 \mapsto 0, \ \overline{z}_0 \mapsto \infty, \ 0 \mapsto e^{i\alpha}\frac{z_0}{\overline{z}_0}?$$

We will now attempt to solve the following.

Problem 3.16. Characterize *all* bilinear transformations that map the upper half-plane onto the interior of the unit circle.

These transformations in question, of course, must map the real line $\mathbb{R}$ onto the unit circle $|w| = 1$. Letting

$$w = \frac{az+b}{cz+d},$$

we choose specific points to determine conditions for our coefficients. Since $|w| = 1$ when $z = 0$ and $z = \infty$, we obtain

$$\left|\frac{b}{d}\right| = 1 \quad \text{and} \quad \left|\frac{a}{c}\right| = 1. \tag{3.13}$$

In view of (3.13), we have

$$\left|\frac{b}{a}\right| = \left|\frac{d}{c}\right|.$$

Hence the transformation may be written as

$$w = \frac{a}{c}\frac{z+b/a}{z+d/c} = e^{i\alpha}\frac{z-z_0}{z-z_1} \qquad \left(z_0 = -\frac{b}{a},\ z_1 = -\frac{d}{c}\right), \tag{3.14}$$

where $\alpha \in \mathbb{R}$ and $|z_0| = |z_1|$. Can we obtain additional information about the relationship between z_0 and z_1? By letting $z = 1$ in (3.14), we have

$$|w| = \left|\frac{1-z_0}{1-z_1}\right| = 1,$$

or

$$|1 - z_0| = |1 - z_1|. \tag{3.15}$$

Upon simplifying (3.15), we see that $\operatorname{Re} z_0 = \operatorname{Re} z_1$. Since $|z_1| = |z_0|$, either $z_1 = z_0$ or $z_1 = \overline{z}_0$. If $z_1 = z_0$, then $ad - bc = 0$ and so, (3.14) reduces to a constant. Thus $z_1 = \overline{z}_0$, and we have

Theorem 3.17. *The most general bilinear transformation of the real line* $\mathbb{R}$ *onto the unit circle* $|w| = 1$ *is given by*

$$w = T(z) = e^{i\alpha}\frac{z-z_0}{z-\overline{z}_0}, \tag{3.16}$$

where $\alpha \in \mathbb{R}$.

Since the point z_0 maps onto the origin and $\overline{z}_0$ maps onto ∞, (3.16) maps the upper half-plane onto the interior of the unit circle if $\operatorname{Im} z_0 > 0$ and onto its exterior if $\operatorname{Im} z_0 < 0$. How do we characterize all bilinear transformations that map the right half-plane $\{z : \operatorname{Re} z > 0\}$ onto the unit disk $|w| < 1$?

We wish to determine the most general set of coefficients such that

$$w = \frac{az+b}{cz+d}$$

will map $\operatorname{Im} z > 0$ onto $\operatorname{Im} w > 0$. As in the previous example, we will first find all bilinear mappings from the boundary of the region in the z plane onto the boundary of the region in the w plane and then determine the subset of these mappings that satisfy our additional criterion. Bilinear mappings from $\operatorname{Im} z = 0$ onto $\operatorname{Im} w = 0$ are found by mapping any three points on the real axis of the z plane onto any three points in the real axis of the w plane. Let $z = z_1, z_2, z_3$ map onto the points $w = 0, 1, \infty$, respectively. Since $w = [0, 1, \infty, w]$, the invariance of cross ratio shows that (see Corollary 3.11)

$$w = (z_1, z_2, z_3, z) = \frac{(z - z_1)(z_2 - z_3)}{(z - z_3)(z_2 - z_1)} \quad \text{for any } z.$$

Since z_1, z_2, z_3 are all real, the coefficients a, b, c, d of

$$w = \frac{az + b}{cz + d}$$

must all be real. To see what further constraints are necessary, we rewrite

$$w = \frac{az + b}{cz + d} = \frac{(az + b)(c\overline{z} + d)}{|cz + d|^2} = \frac{ac|z|^2 + bd + adz + bc\overline{z}}{|cz + d|^2}.$$

Then, whenever $\operatorname{Im} z > 0$, we have

$$\operatorname{Im} w = \frac{(ad - bc)\operatorname{Im} z}{|cz + d|^2} > 0 \ \text{ if and only if } ad - bc > 0.$$

Moreover, $\operatorname{Im} w < 0$ if and only if $ad - bc < 0$. Recall that the map is *onto*. Hence we have the following.

Theorem 3.18. *The most general bilinear map of the upper half-plane* $\{z : \operatorname{Im} z > 0\}$ *onto itself is given by*

$$w = \frac{az + b}{cz + d},$$

where a, b, c, d *are real and* $ad - bc > 0$.

Corollary 3.19. *The most general bilinear map of the upper half-plane* $\{z : \operatorname{Im} z > 0\}$ *onto the lower half-plane* $\{w : \operatorname{Im} w < 0\}$ *is given by*

$$w = \frac{az + b}{cz + d},$$

where a, b, c, d *are real and* $ad - bc < 0$.

Next we ask

Problem 3.20. Determine all bilinear transformations that map the unit disk $\{z : |z| < 1\}$ onto itself.

The answer to this problem will necessitate mapping the unit circle onto itself. Any rotation of the identity function (that is, $w = az$, $|a| = 1$) is a solution to our problem. But there are more general transformations that take the unit disk onto itself.

Theorem 3.21. *The most general bilinear transformation that takes the unit disk $\Delta = \{z : |z| < 1\}$ onto itself is given by*

$$f(z) = e^{i\alpha}\left(\frac{z - z_0}{1 - \overline{z}_0 z}\right), \tag{3.17}$$

where $\alpha \in \mathbb{R}$ and $z_0 \in \Delta$.

If, in Theorem 3.21, we take z_0 with $|z_0| > 1$, then f maps $|z| > 1$ onto $|w| < 1$ so that f in this choice maps $|z| < 1$ onto $|w| > 1$.

There are several proofs of this result. The easiest proof follows from the principle of inverse points. Let us now discuss an important concept concerning inverse/symmetric points. Let L be a line in $\mathbb{C}$. Two points z and z^* are called the *inverse points* (*symmetric*) *with respect to the line* L if L is the perpendicular bisector of $[z, z^*]$, the line segment connecting z and z^*, see Figure 3.5.

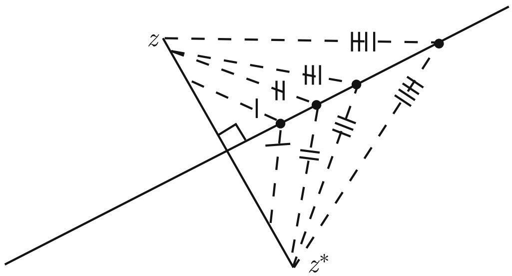


Figure 3.5. Inverse points with respect to a line L

Then it is easy to see that every line or circle passing through both z and z^* intersect L at right angles. For instance,

(i) z and z^* are inverse points with respect to the real axis whenever $z^* = \overline{z}$,
(ii) z and z^* are inverse points with respect to the imaginary axis whenever $z^* = -\overline{z}$.

Consider $w = 1/z$, $z \in \Delta = \{z : |z| < 1\}$. Then the point $z = re^{i\theta}$ $(0 < r < 1)$ in Δ maps onto the point $(1/r)e^{-i\theta}$ which lies outside the unit circle $|z| = 1$.

Let L be the line from the center "O" through $z = re^{i\theta}$. Draw a line S perpendicular to the line L through the point z. The line S intersects the unit circle $|z| = 1$ at two points. Draw tangents at which S intersects the unit circle. It is easy to see these two tangents intersect the line L at the point $1/\overline{z}$. Note that

$$|z|\left|\frac{1}{\overline{z}}\right| = 1$$

and $0, z, 1/\overline{z}$ lie on the same ray.

Using the above discussion, we may define the inverse points with respect to an arbitrary circle as follows: we say that two points z and z^* in $\mathbb{C}$ are the *inverse points with respect to a circle* in $\mathbb{C}$ if every line or circle passing through both z and z^* intersect at right angles. It is easy to see the following: "Let $C = \{\zeta : |\zeta - z_0| = R\}$ be a circle in $\mathbb{C}$ with center at z_0 and radius R. *Two points z and z^* are inverse points with respect to the circle C if*

(i) *z and z^* are collinear with center z_0*
(ii) $|z - z_0|\,|z - z^*| = R^2$."

We remark the following:

- If z moves close to the boundary of C, the point z^* also moves closer to the boundary. In other words, every point on the circle is the inverse point of itself.
- If z moves towards the center z_0, then $|z-z_0| \to 0$ whereas $|z-z^*| \to \infty$. This fact is expressed by saying that the center "z_0" and the point at "∞" are the inverse points with respect to the circle C. Since R is arbitrary, the center and the point ∞ are inverse points with respect to any circle centered at z_0 and any finite radius.
- Let z be a point inside the circle. Then $z = z_0 + re^{i\theta}(r < R)$. If z^* is the inverse point of z with respect to the circle C, then, since z and z^* lie on the same ray through z_0, we have

$$\mathrm{Arg}\,(z^* - z_0) = \mathrm{Arg}\,(z - z_0) = \theta \quad \text{and} \quad |z - z_0|\,|z^* - z_0| = R^2.$$

This gives

$$z^* - z_0 = \left(\frac{R^2}{|z - z_0|}\right)e^{i\theta} = \frac{R^2}{re^{-i\theta}} = \frac{R^2}{\overline{z - z_0}}.$$

The fact discussed above may be formulated as

Corollary 3.22. *Two points z and z^* are inverse points with respect to the circle $C = \{\zeta : |\zeta - z_0| = R\}$ if and only if*

$$(z^* - z_0)(\overline{z - z_0}) = R^2, \quad i.e., \quad z^* = z_0 + \frac{R^2}{\overline{z - z_0}}. \tag{3.18}$$

Thus, one may define the inversion in $C = \{\zeta : |\zeta - z_0| = R\}$ as a map $J_C : \mathbb{C}_\infty \to \mathbb{C}_\infty$ defined by

$$J_C(z) = z_0 + \frac{R^2}{\overline{z - z_0}}, \quad z \neq z_0.$$

Note that $J_C(z_0) = \infty$, $J_c(\infty) = z_0$, and $J_C(\zeta) = \zeta$ whenever $\zeta \in C$. So, the points that are images of each other under J_C are said to be the inverse points with respect to the circle C.

Example 3.23. If α and α^* are inverse points with respect to the circle $|\zeta - z_0| = R$, then we see that the equation of the circle is

$$\left|\frac{\zeta - \alpha}{\zeta - \alpha^*}\right| = k \quad (k \in \mathbb{R}, k \neq 1).$$

To see this, we let $\zeta = z_0 + Re^{i\theta}$. As α and α^* are inverse points with respect to the circle $|\zeta - z_0| = R$, we have

$$\alpha^* = z_0 + \frac{R^2}{r} e^{i\phi}$$

where $\alpha - z_0 = re^{i\phi}$. Then

$$\zeta - \alpha = z_0 + Re^{i\theta} - \alpha = Re^{i\theta} - re^{i\phi} \text{ and } \zeta - \alpha^* = Re^{i\theta} - \frac{R^2}{r} e^{i\phi}$$

so that

$$\left|\frac{\zeta - \alpha}{\zeta - \alpha^*}\right| = \left|\frac{Re^{i\theta} - re^{i\phi}}{Re^{i\theta} - \frac{R^2}{r} e^{i\phi}}\right| = \frac{r}{R}\left|\frac{Re^{i\theta} - re^{i\phi}}{re^{i\theta} - Re^{i\phi}}\right| = \frac{r}{R}$$

and the result follows. ●

Let us now obtain a necessary and sufficient condition for two points z and z^* to be inverse point with respect to a circle in $\mathbb{C}_\infty$.

Let the equation of a line L be

$$aX + bY + c = 0 \quad (a, b, c \in \mathbb{R}),$$

or equivalently in complex form as

$$\overline{\beta} Z + \beta \overline{Z} + c = 0 \quad (\beta = (a + ib)/2 \in \mathbb{C},\ c \in \mathbb{R}).$$

Suppose that $z = x + iy$ and $z^* = x^* + iy^*$ are the inverse points with respect to L. Then the slope of the line passing through z and z^* is

$$m = \frac{b}{a} = \frac{y - y^*}{x - x^*}$$

because the line L has slope $m' = -a/b$ and L is perpendicular to $[z, z^*]$. Thus, as the midpoint $(z + z^*)/2$ of $[z, z^*]$ lies in L, elementary geometry reveals that

$$\frac{y - y^*}{x - x^*} = \frac{b}{a} \text{ and } \overline{\beta}\left(\frac{z + z^*}{2}\right) + \beta\left(\frac{\overline{z} + \overline{z^*}}{2}\right) + c = 0,$$

or equivalently,

$$\overline{\beta}z+\beta\overline{z^*}+c=\beta\overline{z}+\overline{\beta}z^*+c \text{ and } (\overline{\beta}z+\beta\overline{z^*}+c)+(\beta\overline{z}+\overline{\beta}z^*+c)=0.$$

Note that this is of the form $A=\overline{A}$ and $A+\overline{A}=0$ which imply that $A=0$. In conclusion, we have

Theorem 3.24. *Two points z and z^* are inverse points with respect to the line $\overline{\beta}Z+\beta\overline{Z}+c=0$ if and only if $\overline{\beta}z+\beta\overline{z^*}+c=0$.*

Let us state and prove a general result which covers the case of a circle.

Theorem 3.25. *Two points z and z^* are inverse points with respect to the circle in $\mathbb{C}_\infty$,*

$$\alpha Z\overline{Z}+\overline{\beta}Z+\beta\overline{Z}+\gamma=0 \tag{3.19}$$

if and only if

$$\alpha z\overline{z^*}+\overline{\beta}z+\beta\overline{z^*}+\gamma=0. \tag{3.20}$$

(*Note that line is considered as a circle of infinite radius; in this case $\alpha=0$*).

Proof. Without loss of generality, we may assume that $\alpha=1$ as $\alpha=0$ has been dealt with in Theorem 3.24. For $\alpha=1$, (3.19) is equivalent to

$$|Z+\beta|=\sqrt{|\beta^2|-\gamma}.$$

Thus, by (3.18), z and z^* are symmetric with respect to the circle

$$|Z+\beta|=\sqrt{|\beta^2|-\gamma} \quad (z_0=-\beta \text{ and } R=\sqrt{|\beta^2|-\gamma})$$

if and only if

$$[z^*-(-\beta)][\overline{(z-(-\beta))}]=|\beta|^2-\gamma, \quad \text{i.e.,} \quad \overline{z}z^*+\beta\overline{z}+\overline{\beta}z^*+\gamma=0$$

and the proof is complete. ∎

If we choose $\beta=0$ and $\gamma=-1$ in Theorem 3.25, then we see that z and z^* are the inverse points with respect to the unit circle $|z|=1$ if and only if

$$z\overline{z^*}=1, \quad \text{i.e.,} \quad z^*=\frac{1}{\overline{z}}.$$

The concept of inverse points is useful in solving mapping problems that involve bilinear transformations because of the remarkable property which asserts that *"bilinear transformation preserve inverse points"*. More precisely, we have

Theorem 3.26. *Let C be a circle in $\mathbb{C}_\infty$. Suppose further that $w = T(z)$ is a bilinear transformation and $C' = T(C)$ is the transformed circle in $\mathbb{C}_\infty$. The two points z and z^* in $\mathbb{C}_\infty$ are inverse points with respect to C if and only if $w = T(z)$ and $w^* = T(z^*)$ are inverse points with respect to C'.*

Proof. Let the equation of the circle C in $\mathbb{C}_\infty$ be given by

$$\alpha Z\overline{Z} + \overline{\beta}Z + \beta\overline{Z} + \gamma = 0 \quad (\alpha \in \mathbb{R}, \beta \in \mathbb{C}, \gamma \in \mathbb{R}). \tag{3.21}$$

Suppose that z and z^* are a pair of inverse points with respect to the circle C. By Theorem 3.25, z and z^* must satisfy

$$\alpha z\overline{z^*} + \overline{\beta}z + \beta\overline{z^*} + \gamma = 0. \tag{3.22}$$

Let $W = T(Z) = (aZ + b)/(cZ + d)$ be a bilinear transformation. Then

$$Z = \frac{dW - b}{-cW + a}.$$

We wish to show that $T(z) = w$ and $T(z^*) = w^*$ are inverse points with respect to the transformation circle $C' = T(C)$. First we see from (3.21) that the image of the circle C under $W = T(Z)$ is given by

$$\alpha\left(\frac{dW - b}{-cW + a}\right)\left(\frac{\overline{dW} - \overline{b}}{-\overline{c}\overline{W} + \overline{a}}\right) + \overline{\beta}\left(\frac{dW - b}{-cW + a}\right) + \beta\left(\frac{\overline{dW} - \overline{b}}{-\overline{c}\overline{W} + \overline{a}}\right) + \gamma = 0.$$

The image of (3.22) under $w = T(z)$ is the same as above except that w and $\overline{w}$ are replaced by w and $\overline{w^*}$, respectively. Therefore, by Theorem 3.25, w and $\overline{w^*}$ must be inverse points with respect to the transformed circle C' described by the above equation. The converse follows similarly. ∎

Using this theorem, it is easy to characterize all bilinear transformations which map a circle in $\mathbb{C}_\infty$ to another given circle in $\mathbb{C}_\infty$.

Let us first find all bilinear transformations which map the unit disk $\Delta = \{z : |z| < 1\}$ onto itself. To do this let $f(z)$ to be a general bilinear transformation which takes Δ onto itself. Clearly, there exists a z_0 in Δ such that $f(z_0) = 0$. We know that z_0 and $1/\overline{z}_0$ are inverse points with respect to the unit circle $|z| = 1$ (Recall that, as $z_0 \mapsto 0$, $1/\overline{z}_0 \mapsto \infty$ and so 0 and ∞ are inverses with respect to the any circle centered at the origin). As $z_0 \mapsto 0$, $1/\overline{z}_0 \mapsto \infty$, and so f must be of the form

$$w = f(z) = k\left(\frac{z - z_0}{z - 1/\overline{z}_0}\right) = -k\overline{z}_0\left(\frac{z - z_0}{1 - z\overline{z}_0}\right) = A\left(\frac{z - z_0}{1 - z\overline{z}_0}\right)$$

where A is a constant chosen so that $|w| = 1$. As $|z| = 1$ implies that $|w| = 1$, we in particular have

$$|f(1)| = 1, \quad \text{i.e.,} \quad \left|A\left(\frac{1 - z_0}{1 - \overline{z}_0}\right)\right| = |A| = 1,$$

which gives $A = e^{i\alpha}$ for some real α. Thus, f has the desired form, namely (3.17). Theorem 3.21 follows.

Let us next use the symmetry principle to obtain all bilinear transformations that map the upper half-plane $\{z : \operatorname{Im} z > 0\}$ onto the unit disk $|z| < 1$. To see this we suppose that z_0 ($\operatorname{Im} z_0 > 0$) is mapped to $w = 0$. Note that $\overline{z}_0$ is symmetric to z_0 with respect to the real axis of the z-plane. As $w = 0$ and $w = \infty$ are symmetric with respect to the unit circle, the desired bilinear transformation $w = T(z)$ must carry $\overline{z}_0$ to $w = \infty$. Therefore,

$$w = T(z) = A\left(\frac{z - z_0}{z - \overline{z}_0}\right)$$

for some complex constant A. When z is real, we have $|w| = 1$ which gives $|A| = 1$, i.e., $A = e^{i\alpha}$ for some $\alpha \in \mathbb{R}$, and $T(z)$ is of the form (3.16). Thus, we have provided an alternate proof of Theorem 3.17.

Questions 3.27.

1. In choosing three points, why is it often convenient to pick $0, 1$, and ∞?
2. When will the sum of bilinear transformations be a bilinear transformation? The product?
3. What kind of bilinear transformation maps ∞ onto itself?
4. What kind of bilinear transformation maps ∞ onto the origin?
5. How many bilinear transformations map more than two points onto themselves?
6. What is the form of a bilinear transformation which has one fixed point $z_1 \in \mathbb{C}$ and the other fixed point at ∞? How about, in particular, $z_1 = 0$?
7. Is there a bilinear transformation having no fixed point?
8. Is there a bilinear transformation having exactly one fixed point? How about $f(z) = z/(2z + 1)$?
9. What can we say about a transformation which has ∞ as the only fixed point? Is it simply the transformation of the form $f(z) = z + \beta$, for some $\beta \in \mathbb{C}$?
10. Why don't we say that $f(z) = \overline{z}$ is a bilinear transformation? Does it have infinitely many fixed points? Is it true that the only bilinear transformation, having more than two fixed points, is the identity transformation?
11. Does the isogonal map $f(z) = \overline{z}$ preserve cross ratio?
12. Without substituting any points, is there a way to determine whether a region maps inside or outside another region?
13. What can you say about the existence of bilinear transformations from a triangle to a square? Triangle to a circle? Triangle to a straight line? Square to a square?
14. From (3.16) we see that a bilinear mapping from the real line to the unit circle is uniquely determined by finding the point that maps onto the origin and one other point. Does this contradict the fact that three points determine a bilinear transformation?

15. Are there any one-to-one mappings of the extended plane onto itself, other than bilinear transformations?

Exercises 3.28.

1. Find the cross ratio of the four roots of $i^{1/4}$ and $1^{1/4}$.
2. Find a bilinear transformation mapping the points
   (a) $z_1 = 0$, $z_2 = i$, $z_3 = -i$ onto $w_1 = i$, $w_2 = -i$, $w_3 = 0$
   (b) $z_1 = 1$, $z_2 = i$, $z_3 = \infty$ onto $w_1 = i$, $w_2 = -1$, $w_3 = 0$
   (c) $z_1 = 2$, $z_2 = -1$, $z_3 = -i$ onto $w_1 = \infty$, $w_2 = -1$, $w_3 = -i$
   (d) $z_1 = 2$, $z_2 = \infty$, $z_3 = i$ onto $w_1 = \infty$, $w_2 = -1$, $w_3 = 1$.
3. Using the invariance property of the cross-ratio, find a bilinear transformation f in each of the following cases:
   (a) $\{1, i, -1\}$ onto $\{1, 0, i\}$
   (b) $\{\infty, i, 0\}$ onto $\{0, i, \infty\}$
   (c) $\{-i, -2+i, 3i\}$ onto $\{4, 1+3i, -2\}$
   (d) $\{0, 1, \infty\}$ onto $\{-i, 1, i\}$.
4. Under the transformation $w = iz/(z-1)$, find the image of
   (a) the closed unit disk $|z| \leq 1$.
   (b) the closed right half-plane $\operatorname{Re} z \geq 0$.
   (c) the closed upper half-plane $\operatorname{Im} z \geq 0$.
   (d) the open infinite sector $\pi/4 < \operatorname{Arg} z < \pi/2$.
5. Under the transformation $w = (z-1)/(z+1)$, find the image of
   (a) $|z| \leq r < 1$ (b) $|z| \leq r \quad (r > 1)$
   (c) $\operatorname{Im} z > 1$ (d) $\operatorname{Im} z > \operatorname{Re} z$.
6. Find conditions for a bilinear transformation to carry a straight line in the z-plane onto the unit circle $|w| = 1$.
7. Let w be a bilinear transformation from the real line onto the unit circle. If z_1 is mapped onto w_1, show that $\overline{z}_1$ is mapped onto $1/\overline{w}_1$.
8. Let w be a bilinear transformation from the unit circle onto itself. If z_1 is mapped onto w_1, show that $1/\overline{z}_1$ is mapped onto $1/\overline{w}_1$.
9. Prove that the cross ratio of four distinct points is real if and only if the four points lie on a circle or on a straight line.
10. If z_1 and z_2 are distinct fixed points of a bilinear transformation $w = T(z)$, show that the transformation may be expressed as
$$\frac{w - z_1}{w - z_2} = K\frac{z - z_1}{z - z_2},$$
where K is a complex constant.
11. If $z_1 \in \mathbb{C}$ and $z_2 = \infty$ are two fixed points of a bilinear transformation $w = T(z)$, show that the transformation may be expressed as
$$w - z_1 = K(z - z_1),$$
for some complex constant K.

12. Show that a bilinear transformation has either 1, 2 or infinitely many fixed points. Establish conditions for each occurrence.
13. We know that 1 and -1 are fixed points of $f_1(z) = z$ and $g_1(z) = 1/z$. Similarly, i and $-i$ are the fixed points of $f_1(z) = z$ and $g_2(z) = -1/z$. Find a general form of the bilinear transformation which has 1 and -1 as its fixed points.
14. Prove that the bilinear transformation

$$w = T(z) = \frac{(z_1 + z_2)z - 2z_1 z_2}{2z - (z_1 + z_2)} \quad (z_1 \neq z_2)$$

has the fixed points z_1 and z_2, and show that $T(T(z)) = z$.
15. Let

$$w_1 = T_1(z) = \frac{a_1 z + b_1}{c_1 z + d_1} \quad \text{and} \quad w_2 = T_2(z) = \frac{a_2 z + b_2}{c_2 z + d_2}.$$

Prove that $(T_1 \circ T_2)(z) = T_1(T_2(z))$ is also a bilinear transformation.
16. Let $T_1(z), T_2(z)$, and $T_3(z)$ be bilinear transformations. Prove that $T_1(T_2T_3)(z) = (T_1T_2)T_3(z)$.
17. For every bilinear transformation $T_1(z)$, show that there exists a bilinear transformation $T_2(z)$ such that $T_1(T_2(z)) = T_2(T_1(z)) = z$, the identity transformation.
18. Exercises 15, 16 and 17 say that the bilinear transformations form a group under composition. Show that this group is not commutative by finding two bilinear transformations $T_1(z)$ and $T_2(z)$ such that $T_1(T_2(z)) \neq T_2(T_1(z))$.
19. Find all bilinear transformations mapping the imaginary axis onto the unit circle.
20. Find a bilinear transformation f which maps the circle $|z + i| = 1$ onto the real line $\mathbb{R}$.
21. Show that a bilinear transformation that maps the disk $|z| \leq r_1$ onto the disk $|w| \leq r_2$ must be of the form

$$w = \frac{e^{i\alpha} r_1 r_2 (z - z_0)}{r_1^2 - \overline{z}_0 z},$$

where $\alpha \in \mathbb{R}$ and $|z_0| < r_1$.
22. Does the bilinear transformation $w = R(1 + iz)/(1 - iz)$ map the upper half-plane $\{z : \operatorname{Im} z > 0\}$ onto the circle $|w| < R$? What bilinear transformation maps $|z| < R$ onto $\operatorname{Im} w > 0$?
23. Show that the bilinear transformations map open sets onto open sets.
24. Suppose that L is the line passing through the points -1 and i. Are the points $z_1 = 3i$ and $z_2 = 2 + i$ inverse with respect to the line L?
25. Determine whether the following pair of points are inverses with respect to the given line:
    (i) $3i$ and $2 + i$ with respect to the line $z - i\overline{z} - 1 - i = 0$
    (ii) $3i$ and $2 + i$ with respect to the line $z + i\overline{z} + 1 + i = 0$.

26. Determine the inverse point of $1+i$ with respect to the circle $|z+1-2i|=2$.

3.3 Other Mappings

In this section we examine the mapping properties of functions other than bilinear transformations. Consider the function $w=z^2$. Separating this into its real and imaginary parts we obtain

$$w=u+iv=(x+iy)^2=x^2-y^2+i(2xy).$$

This function maps the point (a,a) in the z plane onto the point $(0,2a^2)$ in the w plane. That is, the ray $y=x$, with $x>0$, is mapped onto the ray $(0,v)$, with $v>0$; and the ray $y=x$, $x<0$, is also mapped onto the ray $(0,v)$, $v>0$. In other words, the line $y=x$ is twice mapped onto the ray $(0,v)$, $v\geq 0$ (see Figure 3.6). Observe that, unlike bilinear transformations, the function $w=z^2$ is not one-to-one.

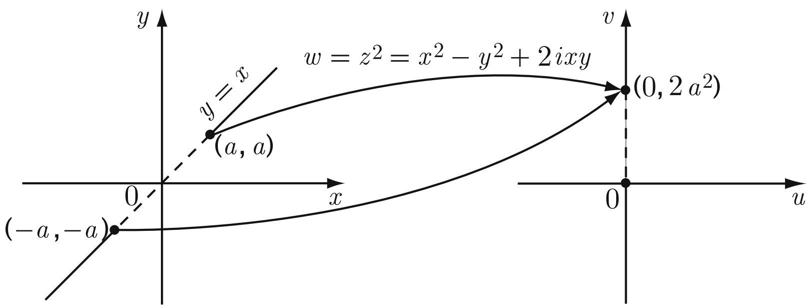


Figure 3.6. Image of the line $y=x$ under $w=z^2$

In general the point $(x,y)=(x,mx)$ is mapped onto the point

$$(u,v)=((1-m^2)x^2,2mx^2).$$

Since

$$\frac{v}{u}=\frac{2m}{1-m^2}\quad (m\neq 1),$$

the straight line $y=mx$ is mapped twice onto the ray

$$v=\left(\frac{2m}{1-m^2}\right)u,$$

where u assumes all the nonnegative real numbers if $|m|<1$ and all nonpositive real numbers if $|m|>1$. Note that the region $0<\text{Arg } z<\pi/4$ is mapped onto the first quadrant, $0<\text{Arg } z<\pi/2$.

To determine the preimage of a circle for the function $w = z^2$, we write

$$u^2 + v^2 = (x^2 - y^2)^2 + (2xy)^2 = (x^2 + y^2)^2.$$

Hence a circle in the z plane, having its center at the origin and radius r, is mapped onto a circle in the w plane having the center at the origin and radius r^2. It is, perhaps, more natural to discuss this function in terms of its polar coordinate representation. We have

$$w = z^2 = [r(\cos\theta + i\sin\theta)]^2 = r^2(\cos 2\theta + i\sin 2\theta).$$

Thus a point with polar coordinates (r, θ) in the z plane is mapped onto the point with polar coordinates $(r^2, 2\theta)$ in the w plane, a point whose distance from the origin is squared and whose argument is doubled. For instance, we have

- the function $f(z) = z^2$ maps the right half-plane
$$\{z : \operatorname{Re} z > 0\} := \{z = re^{i\theta} : 0 < r < \infty, |\theta| < \pi/2\}$$
onto the slit plane $\mathbb{C}\setminus(-\infty, 0]$.
- for each fixed θ_0 with $0 < \theta_0 \leq \pi/2$, the function $f(z) = z^2$ maps the sector $|\text{Arg } z| < \theta_0$ onto the sector $|\text{Arg } w| < 2\theta_0$.

The function $w = z$ and $w = z^2$ both map the unit circle onto itself; but these mappings can no more be considered identical than can the real-valued functions $y = x$ and $y = x^2$, both mapping the closed interval $[0, 1]$ onto itself. The function $w = z^2$ describes the unit circle twice; in fact, it maps any semicircle centered at the origin onto a circle.

We should not leave the function $w = z^2$ without comparing it with its real-valued counterpart, the parabola $y = x^2$. The line $y = c$ in the z plane is transformed into $u = x^2 - c^2$ and $v = 2xc$, from which we obtain

$$u = \left(\frac{v}{2c}\right)^2 - c^2 = \frac{v^2}{4c^2} - c^2.$$

Hence the horizontal line $y = c \neq 0$ is mapped onto the parabola

$$u = \frac{v^2}{4c^2} - c^2.$$

If $c = 0$, the parabola degenerates into the ray $(u, 0), u \geq 0$. In a similar fashion, we can show that the vertical line $x = a \neq 0$ maps onto the parabola (see Figure 3.7)

$$u = -\left(\frac{v^2}{4a^2} - a^2\right).$$

For n a positive integer, the function

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

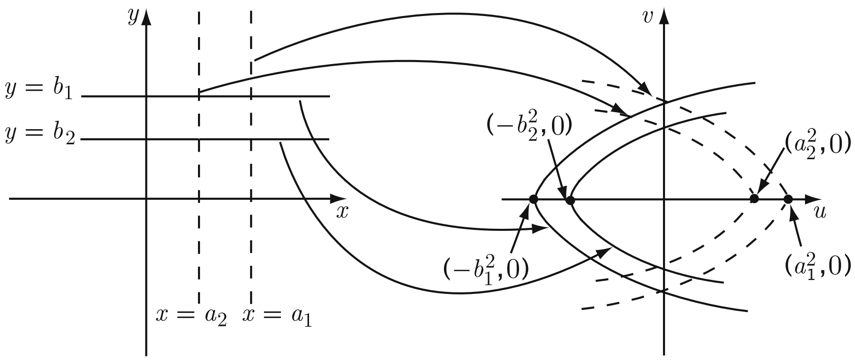


Figure 3.7. Image of lines parallel to coordinate axes under $w = z^2$

maps the point whose distance from the origin is r onto points of distance r^n, and points whose argument is θ onto points having argument $n\theta$. The function $f(z) = z^n$ maps the arc

$$(r, \theta), \quad \theta_0 \le \theta < \theta_0 + (2\pi/n),$$

onto a circle of radius r^n centered at the origin (see Figure 3.8). Moreover, for each fixed θ_0 with $0 < \theta_0 \le \pi/n$, we see that the function $f(z) = z^n$ maps the sector $\{z : |\mathrm{Arg}\, z| < \theta_0\}$ onto the sector $\{w : |\mathrm{Arg}\, w| < n\theta_0\}$. In particular, $f(z) = z^n$ maps the sector $\{z : |\mathrm{Arg}\, z| < \pi/(2n)\}$ onto the half-plane $\{w : \mathrm{Re}\, w > 0\}$.

The function

$$w = \overline{z} = x - iy = r(\cos\theta - i \sin\theta)$$

is yet another function mapping the unit circle onto itself. It maps the point (r, θ) onto the point $(r, -\theta)$. Thus the image of the unit circle described counterclockwise is the unit circle described clockwise (see Figure 3.9). Note that the upper half-plane is mapped onto the lower half-plane.

While the composition of bilinear transformations is again a bilinear transformation (Exercise 3.28(15)), the sum of bilinear transformations need not

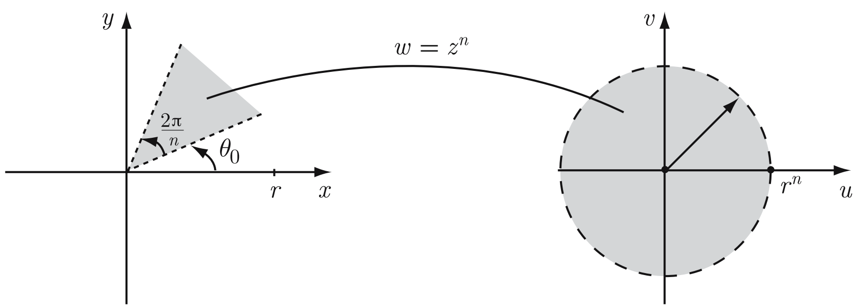


Figure 3.8. Image of an unbounded sector under $w = z^n$

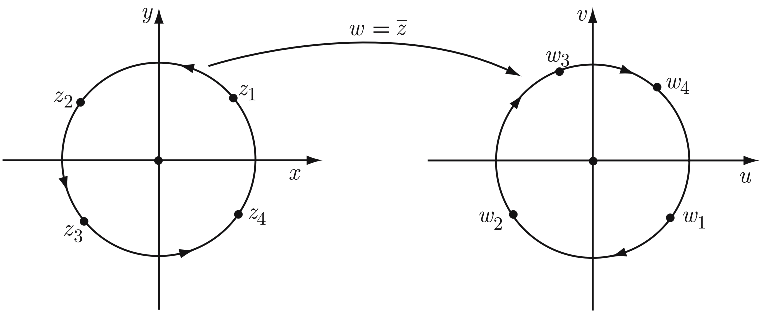


Figure 3.9. Image of a circle under $w = \overline{z}$

be. The last function we will examine in this chapter is

$$w = \frac{1}{2}\left(z + \frac{1}{z}\right).$$

Set $z = re^{i\theta}$. Then, separating this into real and imaginary parts, we have

$$\begin{aligned} w = u + iv &= \frac{1}{2}\left(z + \frac{1}{z}\right) \\ &= \frac{1}{2}\left(r(\cos\theta + i\sin\theta) + \frac{1}{r(\cos\theta + i\sin\theta)}\right) \\ &= \frac{1}{2}\left(r + \frac{1}{r}\right)\cos\theta + i\frac{1}{2}\left(r - \frac{1}{r}\right)\sin\theta. \end{aligned}$$

The unit circle $|z| = 1$ is mapped onto $w = u = \cos\theta$. As θ describes the interval $[0, \pi]$, u decreases continuously from 1 to -1; as θ describes the interval $[\pi, 2\pi]$, u describes the interval $[-1, 1]$. Hence the upper and lower halves of the unit circle are both mapped onto the closed interval $[-1, 1]$.

It is interesting to note that the points z and $1/z$ both map onto the same points under this transformation. Since $1/z$ lies outside the unit circle if and only if z lies inside, it suffices to study the mapping properties for $|z| > 1$.

From the relations

$$u = \frac{1}{2}\left(r + \frac{1}{r}\right)\cos\theta, \quad v = \frac{1}{2}\left(r - \frac{1}{r}\right)\sin\theta, \tag{3.23}$$

we see that, for $r > 1$,

$$\left(\frac{u}{\frac{1}{2}(r + 1/r)}\right)^2 + \left(\frac{v}{\frac{1}{2}(r - 1/r)}\right)^2 = \cos^2\theta + \sin^2\theta = 1.$$

That is, the circle $|z| = r > 1$ is mapped onto an ellipse with major axis along the u axis (Note also that $f(z)$ maps the circle $|z| = 1/r$ onto the same

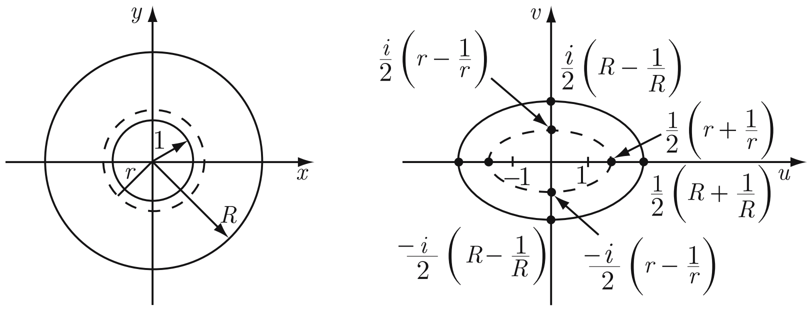


Figure 3.10. Image of circles under the mapping $w = (z + 1/z)/2$

ellipse). As r increases, the ellipse becomes more circular, and as r decreases to 1, the ellipse degenerates to the interval $[-1, 1]$ (see Figure 3.10). We next determine what happens to the ray $\operatorname{Arg} z = \theta$. For $r \geq 1$, we see from (3.23) that the rays $\operatorname{Arg} z = 0$ and $\operatorname{Arg} z = \pi$ are mapped onto themselves, although (excluding the point at ∞) only the points $(1, 0)$ and $(-1, 0)$ remain fixed. Similarly, $\operatorname{Arg} z = \pi/2 \quad (r > 1)$ is mapped onto $\operatorname{Arg} w = \pi/2 \quad (r > 0)$ and $\operatorname{Arg} z = -\pi/2 \quad (r > 1)$ is mapped onto $\operatorname{Arg} w = -\pi/2 \quad (r > 0)$. For all other values of θ we have, according to (3.23),

$$\left(\frac{u}{\cos\theta}\right)^2 - \left(\frac{v}{\cos\theta}\right)^2 = \frac{1}{4}\left[\left(r + \frac{1}{r}\right)^2 - \left(r - \frac{1}{r}\right)^2\right] = 1, \tag{3.24}$$

which is the equation for a hyperbola. For $r > 1$, each arc of this hyperbola is located in the same quadrant as the ray $\operatorname{Arg} z = \theta$ (see Figure 3.11). To summarize, the function

$$w = \frac{1}{2}\left(z + \frac{1}{z}\right)$$

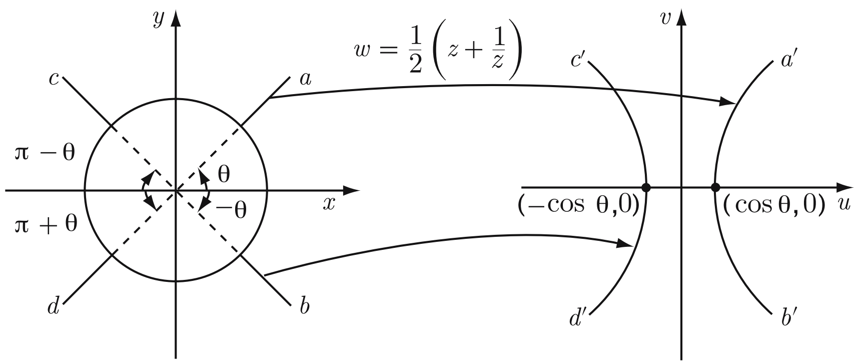


Figure 3.11. Image of lines under the mapping $w = (z + 1/z)/2$

maps the unit circle onto the closed interval $[-1,1]$ twice, and all other circles onto ellipses. It maps both the interior and exterior of the unit circle onto the extended complex plane, excluding the real interval $[-1,1]$. Finally, it maps rays having constant arguments onto arcs of the hyperbolas.

Questions 3.29.

1. For the function $w = f(z) = z^2$, how does the image of $y = c$ differ from that of $y = -c$?
2. What is the largest domain in which the function $w = f(z) = z^2$ is one-to-one?
3. Why was it is more convenient to discuss the function $w = \frac{1}{2}(z + 1/z)$ than $w = z + 1/z$?
4. When does a ray have constant argument?
5. What is the largest domain for which the function $f(z) = \frac{1}{2}(z + 1/z)$ will be one-to-one? Is $f(z)$ one-to-one on the exterior domain $|z| > 1$?
6. How might the mapping properties of the last two sections be combined?

Exercises 3.30.

1. Show that the function $w = z^2$ maps the hyperbolas $x^2 - y^2 = C$ and $xy = K$ onto straight lines.
2. Find the image of the region bounded by straight lines $x = 1$, $y = 1$ and $x + y = 1$ under the mapping. $f(z) = z^2$.
3. Show that $w = ((1+z)/(1-z))^2$ maps the disk $|z| < 1$ onto the plane, excluding the ray $(u, 0)$, $u \leq 0$.
4. Show that $w = z/(1-z)^2$ maps the disk $|z| < 1$ onto the plane, excluding the ray $(u, 0), u \leq -\frac{1}{4}$.
5. Show that the function $w = z^2$ maps the disk $|z - 1| \leq 1$ onto the cardioid $R = 2(1 + \cos\theta)$
6. Discuss the mapping properties of $w = z^{-n}$, n a positive integer.
7. Find a transformation which maps $\Omega_n = \{z : 0 < \operatorname{Arg} z < \pi/n\}$ $(n \in \mathbb{N})$ onto the unit disk $|w| < 1$.
8. Find the image of the sector $|z| < 1$, $0 < \operatorname{Arg} z < \pi/n$, for the function
   (a) $w = \dfrac{z^n + 1}{z^n - 1}$ (b) $w = \left(\dfrac{z^n + 1}{z^n - 1}\right)^2$.
9. Find the image of the unit disk $|z| \leq 1$ for the function
$$w = f(z) = \prod_{k=1}^{n} \frac{|z_k|}{z_k} \left(\frac{z_k - z}{1 - \overline{z}_k z}\right),$$
   where $0 < |z_k| < 1$ for every k.

4

Elementary Functions

Many high school students are puzzled by the following "proof": Let $a = b$. Then

$$a^2 = ab, \; a^2 - b^2 = ab - b^2, \quad \text{and} \quad (a+b)(a-b) = b(a-b).$$

Dividing by $a - b$, we have

$$a + b = b, \; 2b = b, \quad \text{and} \quad 2 = 1.$$

The reader, of course, is not fooled by the invalid division by zero. So let us produce an absurdity without dividing by zero. Since $1/(-1) = (-1)/1$, we take square roots to obtain

$$\sqrt{(1/-1)} = \sqrt{(-1/1)}, \; \sqrt{1}/\sqrt{-1} = \sqrt{-1}/\sqrt{1}, \quad \text{and} \quad 1/i = i/1.$$

Cross multiplying, we have $1^2 = i^2$ or $1 = -1$.

In this chapter, we will show that 1 does not really equal -1. We will also see that the complex exponential and trigonometric functions have much in common, and that a function having a complex exponent must be defined in terms of a logarithm.

4.1 The Exponential Function

Recall that the real-valued function $f(x) = e^x$ has the following properties:

1. e^x is continuous on $\mathbb{R}$, $e^x > 0$, $e^{-x} = 1/e^x > 0$,
2. $e^x \to +\infty$ as $x \to +\infty$, $e^{-x} \to 0$ as $x \to +\infty$,
3. e^x is equal to its derivative,
4. e^x has the power series expansion

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad \text{for } x \in \mathbb{R},$$

5. the rule of exponents,

$$e^{x_1}e^{x_2} = e^{x_1+x_2} \text{ for } x_1, x_2 \in \mathbb{R}.$$

The function $f(x) = e^x$ maps the set of real numbers *one-to-one onto* the set of positive reals. Therefore, it has a continuously strictly increasing inverse function called *natural logarithm*:

$$\ln : \mathbb{R}^+ \to \mathbb{R}; \quad \text{i.e.,} \quad x = \ln y \Longleftrightarrow y = e^x.$$

Now we raise the following.

Problem 4.1. Can we extend the definition of the real exponential function to the complex case? If so, what properties remain the same?

In defining the complex-valued function

$$w = f(z) = e^z = e^{x+iy},$$

we would like to preserve the important properties of the corresponding real-valued function. If the rule of the exponents is to hold, we must have

$$e^z = e^{x+iy} = e^x e^{iy}.$$

It remains to give a "reasonable" definition for e^{iy}.

If we could expand e^{iy} in a power series similar to that of e^x, we would have

$$e^{iy} = 1 + iy + \frac{(iy)^2}{2!} + \frac{(iy)^3}{3!} + \frac{(iy)^4}{4!} + \frac{(iy)^5}{5!} + \cdots . \tag{4.1}$$

Separating (4.1) into its real and imaginary parts, we would obtain

$$e^{iy} = \left(1 - \frac{y^2}{2!} + \frac{y^4}{4!} + \cdots\right) + i\left(y - \frac{y^3}{3!} + \frac{y^5}{5!} + \cdots\right). \tag{4.2}$$

The power series expansion in (4.2) represents the functions $\cos y$ and $\sin y$, respectively. This leads to the following *definition*:

$$e^{iy} = \cos y + i \sin y \quad (y \text{ real}). \tag{4.3}$$

We emphasize that (4.3) is a definition, and that the above argument was introduced only to make this definition seem plausible. In Chapter 8, we will formally prove the validity of the complex power series expansion, thus justifying our definition. The familiar De Moivre law,

$$(\cos y + i \sin y)^n = \cos ny + i \sin ny,$$

may now be expressed as $(e^{iy})^n = e^{iny}$. Note that

$$|e^{iy}| = \sqrt{\cos^2 y + \sin^2 y} = 1$$

for any real number y.

Remark 4.2. Setting $y = \pi$ in (4.3), we obtain

$$e^{\pi i} + 1 = 0,$$

which is, in the authors' opinion, the most beautiful equation in all of mathematics. It contains the five most important constants as well as the three most important operations (addition, multiplication and exponentiation). ●

We now examine some of the consequences of defining e^z to be the complex number

$$e^z = e^{x+iy} = e^x(\cos y + i \sin y) \tag{4.4}$$

so that for $z = x$ ($x \in \mathbb{R}$), the definition of e^z coincides with the usual exponential function e^x. For $z = 0 + iy$ ($y \in \mathbb{R}$), the definition agrees with (4.3). If $x = 0$, we have, for any real numbers y_1 and y_2, the addition formula for e^{iy}:

$$\begin{aligned} e^{iy_1}e^{iy_2} &= (\cos y_1 + i \sin y_1)(\cos y_2 + i \sin y_2) \\ &= (\cos y_1 \cos y_2 - \sin y_1 \sin y_2) + i(\cos y_1 \sin y_2 + \sin y_1 \cos y_2) \\ &= \cos(y_1 + y_2) + i \sin(y_1 + y_2) = e^{i(y_1+y_2)}. \end{aligned}$$

Using this, one can obtain the fundamental property of the exponential function, namely the addition formula,

$$e^{z_1}e^{z_2} = e^{z_1+z_2}.$$

To see this, let $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$. Then it follows that

$$\begin{aligned} e^{z_1}e^{z_2} &= e^{x_1+iy_1}e^{x_2+iy_2} \\ &= e^{x_1}e^{iy_1}e^{x_2}e^{iy_2} \\ &= (e^{x_1}e^{x_2})(e^{iy_1}e^{iy_2}) \\ &= e^{x_1+x_2}e^{i(y_1+y_2)} \\ &= e^{(x_1+iy_1)+(x_2+iy_2)} \\ &= e^{z_1+z_2}, \end{aligned}$$

and the rule for exponents remains valid for complex numbers. Similarly,

$$\frac{e^{z_1}}{e^{z_2}} = e^{z_1-z_2}, \quad (e^{z_1})^n = e^{nz_1} \quad \text{for } n \in \mathbb{Z}$$

and we can get De Moivre's formula

$$e^{in\theta} = (e^{i\theta})^n, \qquad \text{i.e., } \cos n\theta + i \ \sin n\theta = (\cos\theta + i \sin\theta)^n, \quad \text{for } n \in \mathbb{Z}.$$

Since $|e^z| = |e^x e^{iy}| = e^x|e^{iy}| = e^x$, we see that $e^z \neq 0$ for any complex number z. Moreover from the addition formula for e^z, we have

$$e^z e^{-z} = e^0 = 1.$$

Consequently, the inverse of e^z is e^{-z}.

Thus, most of the important properties of e^x are preserved for e^z. There is, however, a notable exception. The function e^z is not one-to-one. In fact for any complex number z,

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z(\cos 2\pi + i \sin 2\pi) = e^z.$$

Next, suppose $e^z = e^{x+iy} = 1$. Then,

$$e^x \cos y = 1 \quad \text{and} \quad e^x \sin y = 0.$$

Since $e^x \neq 0$, the second relation gives

$$\sin y = 0; \quad \text{that is, } y = n\pi, \quad n \in \mathbb{Z}.$$

But if we substitute $y = n\pi$ into the first equation, we get

$$e^x \cos n\pi = 1,$$

so that n must be an even integer and in this case, x must be equal to 0. Hence, z is an integral multiple of $2\pi i$. That is,

$$e^z = 1 \Longleftrightarrow z = 2k\pi i, \quad k \in \mathbb{Z}.$$

To summarize the above discussion we need to introduce the definition of periodic function. A function $f : \mathbb{C} \to \mathbb{C}$ is called *periodic* if there exists a complex number such that $f(z + \omega) = f(z)$ for all $z \in \mathbb{C}$. The number ω is then called a *"period"* of f.

Theorem 4.3. *The exponential function $f(z) = e^z$ is periodic with the pure imaginary period $2\pi i$. That is, $e^{z+2\pi i} = e^z$ for all $z \in \mathbb{C}$.*

We have $e^{z+2k\pi i} = e^z$ for $k \in \mathbb{Z}$. In view of this, for any two complex numbers z_1 and z_2 for which $e^{z_1} = e^{z_2}$, we have $e^{z_1 - z_2} = 1$. Consequently, $z_1 - z_2 = 2k\pi i$, k is an integer. Hence, we have

Theorem 4.4. *The equality $e^{z_1} = e^{z_2}$, for $z_1, z_2 \in \mathbb{C}$, holds if and only if $z_1 = z_2 + 2k\pi i$ for some $k \in \mathbb{Z}$.*

If the exponential function e^z assumes a value once, it must—by its periodicity—assume the value infinitely many times. We now show that e^z assumes every finite, nonzero complex number infinitely often. If $e^z = a + ib$, a and b both not 0, then

$$e^x \cos y = a, \quad e^x \sin y = b. \tag{4.5}$$

Squaring both terms in (4.5), we obtain

$$e^{2x}(\cos^2 y + \sin^2 y) = e^{2x} = a^2 + b^2.$$

Since the logarithm is well defined for positive real numbers, we have

$$x = \frac{1}{2}\ln(a^2 + b^2).$$

When $a \neq 0$, we divide the second expression by the first in (4.5) to obtain

$$\tan y = \frac{b}{a}, \quad \text{i.e., } y = \tan^{-1}\left(\frac{b}{a}\right).$$

Hence

$$z = \frac{1}{2}\ln(a^2 + b^2) + i\tan^{-1}\left(\frac{b}{a}\right)$$

is a solution to the equation $e^z = a + bi$. If one of the values of $\tan^{-1}(b/a)$ is y_0, then $y_0 + 2k\pi$, for any integer k, must also be a value. If $a = 0$, from $e^z = 0 + ib = ib$, it follows that

$$e^x = |b|, \quad y = \arg(e^z) = \arg(ib),$$

and therefore

$$z = \begin{cases} \ln|b| + i\left(\frac{\pi}{2} + 2k\pi\right) & \text{if } b > 0 \\ \ln|b| + i\left(-\frac{\pi}{2} + 2k\pi\right) & \text{if } b < 0 \end{cases}, \quad k \in \mathbb{Z}.$$

Example 4.5. Let $e^z = 5 - 5i$. Then,

$$\begin{aligned} z &= \frac{1}{2}\ln[5^2 + (-5)^2] + i\tan^{-1}\left(\frac{-5}{5}\right) \\ &= \frac{\ln 50}{2} + i\left(-\frac{\pi}{4} + 2k\pi\right), \quad k \in \mathbb{Z}. \end{aligned}$$

Suppose $e^z = -5 + 5i$. Then,

$$\begin{aligned} z &= \frac{1}{2}\ln[(-5)^2 + 5^2] + i\tan^{-1}\left(\frac{5}{-5}\right) \\ &= \frac{\ln 50}{2} + i\left(\frac{3\pi}{4} + 2k\pi\right), \quad k \in \mathbb{N}. \end{aligned}$$

Note that

$$\tan^{-1}\left(\frac{-b}{a}\right) \neq \tan^{-1}\left(\frac{b}{-a}\right).$$

●

The definition of the exponential in terms of the trigonometric functions suggests that the process may be reversed. From (4.3) we obtain

$$e^{-iy} = \cos(-y) + i\sin(-y) = \cos y - i\sin y. \tag{4.6}$$

Subtracting or adding (4.3) and (4.6) leads to

$$\sin y = \frac{e^{iy} - e^{-iy}}{2i}, \quad \cos y = \frac{e^{iy} + e^{-iy}}{2}.$$

It therefore seems natural to *define* the complex trigonometric functions by

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}, \quad \cos z = \frac{e^{iz} + e^{-iz}}{2}. \tag{4.7}$$

Replacing z by $-z$ shows that

$$\sin z = -\sin(-z), \quad \cos z = \cos(-z).$$

Note also that $\overline{e^z} = e^{\overline{z}}$. Having extended the definition of $\sin x$ and $\cos x$, it is of interest to note that these complex trigonometric functions do not have any additional zeros. More precisely, we see that

$$\sin z = 0 \iff z = k\pi \ \text{ for some } k \in \mathbb{Z}.$$

Indeed

$$\begin{aligned}
\sin z = 0 &\iff \frac{e^{iz} - e^{-iz}}{2i} = 0 \\
&\iff e^{iz} = e^{-iz} \\
&\iff iz - (-iz) = 2k\pi i \ \text{ for some } k \in \mathbb{Z} \\
&\iff z = k\pi \quad \text{for some} \quad k \in \mathbb{Z},
\end{aligned}$$

as claimed. A similar argument shows that

$$\cos z = 0 \iff z = \left(k + \frac{1}{2}\right)\pi \ \text{ for some } k \in \mathbb{Z}.$$

Also, this result follows from the former, because $\cos z = \sin(z + \pi/2)$. In fact, as

$$e^{i\left(z+\frac{\pi}{2}\right)} = e^{iz}e^{i\pi/2} = ie^{iz} \ \text{ and } \ e^{-i\left(z+\frac{\pi}{2}\right)} = -ie^{-iz},$$

(4.7) gives

$$\sin\left(z + \frac{\pi}{2}\right) = \frac{e^{i(z+\pi/2)} - e^{-i(z+\pi/2)}}{2i} = \frac{ie^{iz} - (-ie^{-iz})}{2i} = \frac{e^{iz} + e^{-iz}}{2}$$

so that, replacing z by $z - \frac{\pi}{2}$ and z by $-z$, respectively, gives

$$\sin z = \cos\left(z - \frac{\pi}{2}\right) \quad \text{and} \quad \sin\left(\frac{\pi}{2} - z\right) = \cos z.$$

Similarly, we have $\sin 2z = 2\sin z \cos z$ because

$$2\sin z\cos z = 2\left(\frac{e^{iz}-e^{-iz}}{2i}\right)\left(\frac{e^{iz}+e^{-iz}}{2}\right) = \frac{e^{2iz}-e^{-2iz}}{2i} = \sin 2z.$$

The remaining trigonometric functions are defined by the usual relations

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

With these definitions, most of the familiar real-valued trigonometric properties can be extended to the complex plane. Now for the exception. The real-valued sine and cosine functions are bounded by 1. However, neither $\sin z$ nor $\cos z$ is bounded in the complex plane. From the triangle inequality, we have

$$|\sin z| = \left|\frac{e^{iz}-e^{-iz}}{2i}\right| \geq \frac{\big||e^{-iz}|-|e^{iz}|\big|}{2} = \frac{|e^{y}-e^{-y}|}{2}.$$

As z approaches infinity along the ray $\operatorname{Arg} z = \pi/2$ or $\operatorname{Arg} z = -\pi/2$, the expression on the right grows arbitrarily large, showing that $\sin z$ is unbounded. Similarly,

$$|\cos z| = \left|\frac{e^{iz}+e^{-iz}}{2}\right| \geq \frac{|e^{y}-e^{-y}|}{2},$$

and $|\cos z|$ also approaches ∞ as z approaches ∞ along the ray $\operatorname{Arg} z = \pm\pi/2$. In fact, to show that $\sin z$ and $\cos z$ are not bounded in $\mathbb{C}$, it suffices to observe that

$$\sin(iy) = \frac{e^{-y}-e^{y}}{2i} \quad \text{and} \quad \cos(iy) = \frac{e^{-y}+e^{y}}{2}$$

showing that each of $|\sin(iy)|$ and $\cos(iy)$ is large whenever y is large.

The identities in (4.7) may be used to find solutions for equations involving the trigonometric functions.

Example 4.6. Let us find all the complex numbers for which $\cos z = 2$. To do this, by the definition of $\cos z$, we must have

$$(e^{iz}+e^{-iz})/2 = 2,$$

which leads to $e^{2iz} - 4e^{iz} + 1 = 0$, a quadratic in e^{iz}. Solving this for e^{iz}, we obtain

$$e^{iz} = \frac{4 \pm \sqrt{16-4}}{2} = 2 \pm \sqrt{3}.$$

But $e^{iz} = e^{i(x+iy)} = e^{-y}(\cos x + i\sin x) = 2 \pm \sqrt{3}$, which gives the relations

$$e^{-y}\cos x = 2 \pm \sqrt{3}, \quad e^{-y}\sin x = 0.$$

The second relation shows that $x = n\pi$, $n \in \mathbb{Z}$; and the first shows that n must be even, so that the last relation reduces to

$$e^{-y} = 2 \pm \sqrt{3}.$$

This gives $y = -\ln(2 \pm \sqrt{3})$. Hence $\cos z = 2$ if and only if

$$z = 2k\pi - i\ln(2 \pm \sqrt{3}) = 2k\pi \pm i\ln(2 + \sqrt{3}), \quad k \in \mathbb{Z}.$$

We leave it as an exercise for the reader to show that both $\sin z$ and $\cos z$ assume every value in the complex plane. ●

Finally, as in the real case, we define the hyperbolic sine and hyperbolic cosine functions by the formulas,

$$\sinh z = \frac{e^z - e^{-z}}{2}, \quad \cosh z = \frac{e^z + e^{-z}}{2}, \quad z \in \mathbb{C}. \tag{4.8}$$

As an immediate consequence of (4.8), we have the relations

$$\sinh z = -i \sin iz, \quad \cosh z = \cos iz.$$

Observe that

$$\sinh(-z) = -\sinh z \quad \text{and} \quad \cosh(-z) = \cosh z.$$

Note also that both $\sinh z$ and $\cosh z$ are periodic, with period $2\pi i$. We may also define

$$\operatorname{sech} z = \frac{1}{\cosh z}, \quad \operatorname{csch} z = \frac{1}{\sinh z}, \quad \tanh z = \frac{\sinh z}{\cosh z}, \quad \coth z = \frac{1}{\tanh z}.$$

Questions 4.7.

1. For what functions $f(z)$ will $e^{f(z)}$ be periodic?
2. What is the largest region in which e^z is one-to-one?
3. What is the largest region in which e^z is bounded?
4. What is the largest region in which $\sin z$ is bounded?
5. When does $e^{f(z)} = e^{\overline{f(z)}}$? Does $\sin(i\bar{z}) = \overline{\sin z}$? Does $\cos(i\bar{z}) = \overline{\cos z}$?
6. When does $e^{f(z)} = e^{\overline{f(z)}}$? Does $\sin(i\bar{z}) = \overline{\sin z}$? Does $\cos(i\bar{z}) = \overline{\cos z}$?
7. When does $\cos z_1 = \cos z_2$? When does $\cos z_1 + \cos z_2 = 0$? When does $\sin z_1 = \sin z_2$? When does $\sin z_1 + \sin z_2 = 0$? When does $e^{z_1} + e^{z_2} = 0$?
8. Are there any trigonometric identities, valid for real variables, that are not valid in the complex plane?
9. How do $|\sin z|$ and $\sin |z|$ compare?
10. How do $|\sin z|$ and $|\sinh z|$ compare?
11. What happens to e^z as $z \to \infty$ along different rays? What about $e^z + z$?
12. Are the zero sets of $\sin z$ in $\mathbb{C}$, and $\sin x$ in $\mathbb{R}$ the same?
13. Are the zero sets of $\cos z$ in $\mathbb{C}$, and $\cos x$ in $\mathbb{R}$ the same?
14. Does the equation $\tan z = i$ have a solution in $\mathbb{C}$?
15. For what values of z is $|\sin z| \le 1$?

Exercises 4.8.

1. Find all values of z for which
   (a) $e^{3z} = 1$ (b) $e^{z^2} = 1$ (c) $e^{e^z} = 1$.
2. Show that all the zeros of $\sin z$ and $\cos z$ are real.
3. (a) Show that both $\sin z$ and $\cos z$ are unbounded on the ray $\operatorname{Arg} z = \theta$, $0 < |\theta| < \pi$.
   (b) Show that $\sin z$ is bounded only on sets contained in a horizontal strip.
4. For $|z| = r$, prove that
   (a) $e^{-r} \le |e^z| \le e^r$
   (b) $e^{-rn} \le |e^{zn}| \le e^{rn}$, n a positive integer.

   When will equality hold?
5. Prove the following identities:
   (a) $\sin^2 z + \cos^2 z = 1$
   (b) $\sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2$
   (c) $\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2$.

   Note: From (a) it appears that both $\sin z$ and $\cos z$ are bounded. But we have already shown that this is not the case!
6. (a) Separate $e^{1/z}$, $z \neq 0$ into its real and imaginary parts.
   (b) Show that $|e^{1/z}|$ is bounded in the region $|z| \ge \epsilon$, $\epsilon > 0$.
7. (a) Prove that e^{iz} is periodic, with period 2π.
   (b) For an arbitrary nonzero complex number a, show that e^{az} is periodic, and find its period.
8. Prove the following inequalities:
   (a) $|e^z + e^{z^2}| \le e^x + e^{x^2-y^2}$
   (b) $|e^{iz} + e^{iz^2}| \le e^{-y} + e^{-2xy}$
   (c) $|\sin z|^2 + |\cos z|^2 \ge 1$.
9. Prove the following hyperbolic identities:[1]
   (a) $\cosh^2 z - \sinh^2 z = 1$
   (b) $\sinh(z_1 \pm z_2) = \sinh z_1 \cosh z_2 \pm \cosh z_1 \sinh z_2$
   (c) $\cosh(z_1 \pm z_2) = \cosh z_1 \cosh z_2 \pm \sinh z_1 \sinh z_2$
   (d) $\sinh z = \sinh x \cos y + i \cosh x \sin y$
   (e) $\cosh z = \cosh x \cos y + i \sinh x \sin y$
   (f) $|\sinh z|^2 = \sinh^2 x + \cos^2 y$
   (g) $|\cosh z|^2 = \sinh^2 x + \cos^2 y$.
10. Show that
    (a) $|\sin z|^2 = \sin^2 x + \sinh^2 y$
    (b) $|\cos z|^2 = \cos^2 x + \sinh^2 y$.
11. Prove that $\tanh z = (\sinh z)/(\cosh z)$ is periodic, with period πi.

[1] The hyperbolic identity (a) clarifies somewhat the adjective “hyperbolic” if one recalls that $x^2 - y^2 = 1$ is the equation of a hyperbola in $\mathbb{R}^2$.

4.2 Mapping Properties

The real and imaginary parts of the nonzero complex number $z = x + iy$ are "equally" important in determining its position in the plane. We have

$$|z| = \sqrt{x^2 + y^2} \quad \text{and} \quad \tan(\arg z) = \frac{y}{x}.$$

The real and imaginary parts of $z = x + iy$ play independent roles, however, in determining the position of the point e^z in the w plane. Separating the function $w = e^z$ into its real and imaginary components, we have

$$w = e^z = u + iv = e^x e^{iy},$$

from which we obtain

$$|e^z| = e^x, \quad \tan(\arg(e^z)) = \frac{v}{u} = \tan y.$$

These relations show that the modulus of e^z depends only on the real part of z, while the argument of e^z depends only on the imaginary part of z. Indeed, as $e^z = e^x e^{iy}$, one has

$$\arg(e^z) = y + 2k\pi, \; k \in \mathbb{Z}.$$

It will therefore be of some interest to determine the image of the lines parallel to our coordinate axes, but first we will make use of the periodicity of the exponential. We have already seen that

$$e^{2k\pi i} = 1, \quad e^z e^w = e^{z+w} \text{ and } e^{z+2k\pi i} = e^z \text{ for every } z \in \mathbb{C} \text{ and } k \in \mathbb{Z},$$

so that the points $x_0 + i(y_0 + 2k\pi)$ have the same image for every integer k. Hence we may examine the mapping properties by restricting ourselves to the infinite strip $-\pi < \operatorname{Im} z \le \pi$. Whatever occurs in this strip will also occur in the strip $-\pi + 2k\pi < \operatorname{Im} z \le \pi + 2k\pi$. With this restriction, $\operatorname{Arg}(e^z) = y$, $-\pi < y \le \pi$. We have the following:

- Since e^z has constant modulus, all the points on the line $x = x_0$ are mapped onto the points equidistant from the origin. In particular, the line segment $x = x_0$, $-\pi < y \le \pi$, is mapped one-to-one onto the circle in the w plane having center at the origin and radius e^{x_0}. As y increases from $-\pi$ to π, the circle is described in a counterclockwise direction.
- Since $|e^z| = e^x > 1$ if and only if $x > 0$, the semi-infinite-strip $\{z : \operatorname{Re} z > 0, -\pi < \operatorname{Im} z \le \pi\}$ is mapped one-to-one onto $\{w : |w| > 1\}$, while the strip $\{z : \operatorname{Re} z < 0, -\pi < \operatorname{Im} z \le \pi\}$ is mapped onto the punctured unit disk $\{w : 0 < |w| < 1\}$ (see Figure 4.1).
- As $|e^z| = e^x < 1$ if and only if $x < 0$, the semi-infinite strip

  $$\{z : \operatorname{Re} z < 0, 0 \le \operatorname{Im} z \le \pi\}$$

  is mapped one-to-one onto the upper semi-disk $\{w : \operatorname{Im} w \ge 0, |w| < 1\}$ excluding the origin.

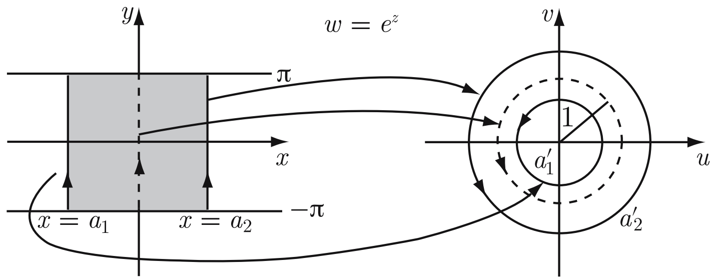


Figure 4.1. Image of line segments parallel to coordinate axes under e^z

As already noted, $\operatorname{Re} z$ plays no role in determining the argument of e^z. Hence the points with identical imaginary parts will map onto the points having the same argument. For the line $y = y_0$, $-\pi < y_0 \leq \pi$, we have

$$w = e^z = e^{x+iy_0} = e^x(\cos y_0 + i \sin y_0).$$

- Since e^x describes the positive reals, the line $y = y_0$ is mapped one-to-one onto the ray $\operatorname{Arg} w = y_0$. Therefore, the infinite strip

  $$\{z : 0 < \operatorname{Im} z < \pi\}$$

  is mapped one-to-one onto the upper half-plane $\{z : \operatorname{Im} z > 0\}$, while the strip

  $$\{z : -\pi < \operatorname{Im} z < 0\}$$

  is mapped onto the lower half-plane $\{z : \operatorname{Im} z < 0\}$ (see Figure 4.2).

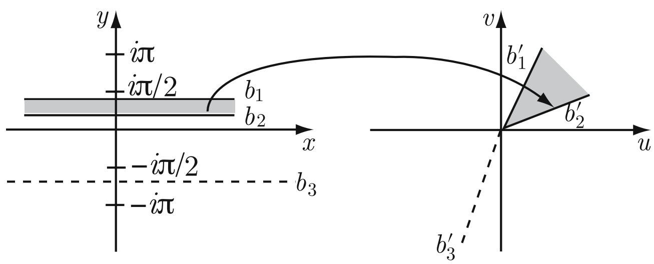


Figure 4.2. Image of lines parallel to the real axis under e^z

- Note that the x axis, $y = 0$, is mapped onto the positive real axis and the line $y = \pi$ is mapped onto the negative real axis. Hence, under the exponential function e^z, the strip

  $$\{z : -\pi < \operatorname{Im} z \leq \pi\}$$

  is mapped one-to-one onto the punctured w plane, $\mathbb{C} \setminus \{0\}$.

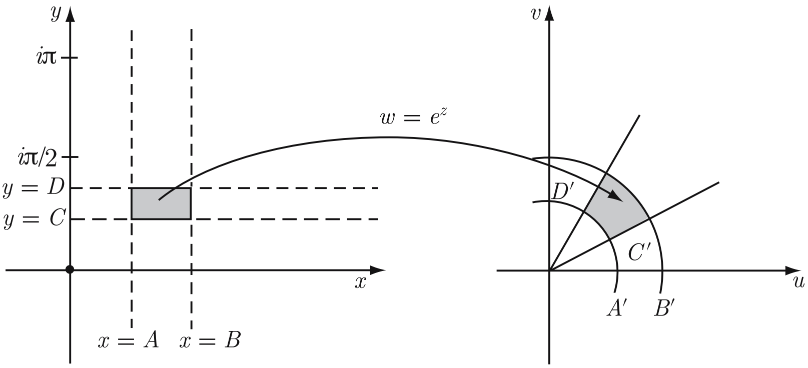


Figure 4.3. Image of a rectangle under e^z

We can combine the two previous mappings to determine the image of the rectangles for the function $w = e^z$. Writing the image in the polar form, we have the rectangle

$$\{z :\ A \le x \le B,\ -\pi < C \le y \le D \le \pi\}$$

being mapped onto the region

$$\{Re^{i\theta} :\ e^A \le R \le e^B,\ C \le \theta \le D\},$$

bounded by arcs and rays (see Figure 4.3).

Next consider a straight line not parallel to either of the coordinate axes. The image of this line will have neither constant modulus nor constant argument, yet it must grow arbitrarily large as x grows arbitrarily large, and must make a complete revolution each time y increases by 2π, thus producing a spiraling effect. If $y = mx + b$, $m \neq 0$, then

$$w = e^z = e^{x+i(mx+b)}.$$

Hence $|e^z| = e^x$ and $\arg(e^z) = mx + b + 2k\pi$, k an integer. In polar form, we may write $w = Re^{i\theta}$, with

$$\begin{cases} R = |e^z| = e^x \\ \theta = \mathrm{Arg}\,(e^z) = mx + b + 2k\pi, \end{cases} \tag{4.9}$$

where $k = k(x)$ is an integer chosen so that θ always satisfies the inequality $-\pi < \theta \le \pi$. Since x describes the set of real numbers, k must describe the set of integers. Eliminating x from the relations in (4.9), we obtain

$$R = e^{(\theta - b - 2k\pi)/m} = e^{-b/m} e^{(\theta - 2k\pi)/m}. \tag{4.10}$$

Letting $\alpha = \theta - 2k\pi$ in (4.10), we have

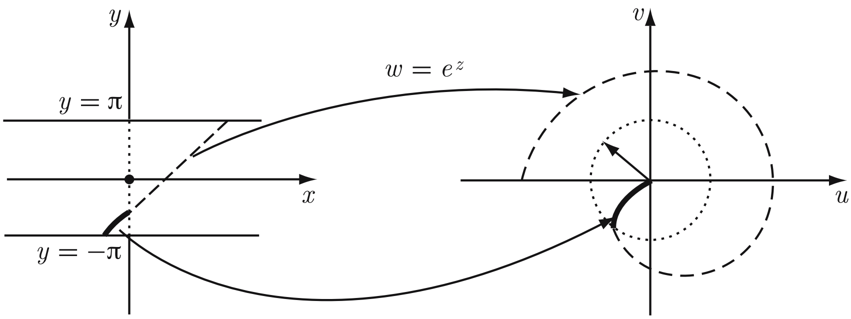


Figure 4.4. Logarithmic spiral

$$R = Ke^{\alpha/m}, \tag{4.11}$$

where K is a positive constant and α describes the set of real numbers.

Equation (4.11) represents what is known as a *logarithmic spiral.* In Figure 4.4 we show the image of one segment of a line, and in Figure 4.5 we show a more complete picture.

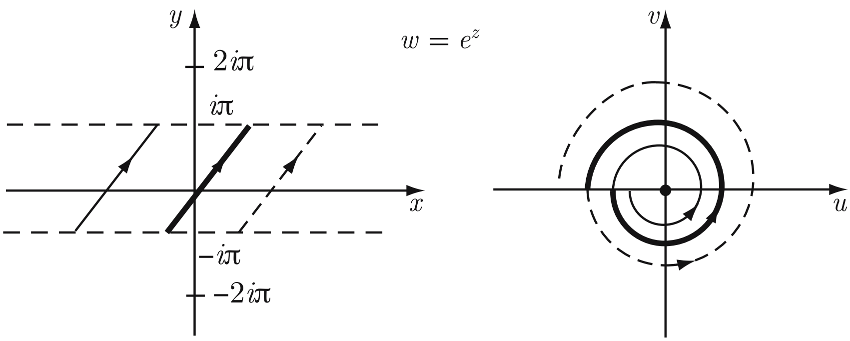


Figure 4.5.

Since the argument of iz and the argument of z differ by $\pi/2$, we expect the function $w = e^{iz}$ to maps lines parallel to the y axis (x axis) onto the same kind of figure as the function $w = e^z$ maps lines parallel to the x axis (y axis). Setting

$$w = e^{iz} = e^{i(x+iy)} = e^{-y+ix},$$

we see that

$$|e^{iz}| = e^{-y} \quad \text{and} \quad \arg(e^{iz}) = x + 2k\pi.$$

Hence the line segment $-\pi < x \leq \pi$, $y = y_0$, is mapped onto the circle having center at the origin and radius e^{-y_0}. The semi-infinite strip,

$$\{z : -\pi < x \leq \pi,\ y > 0\},$$

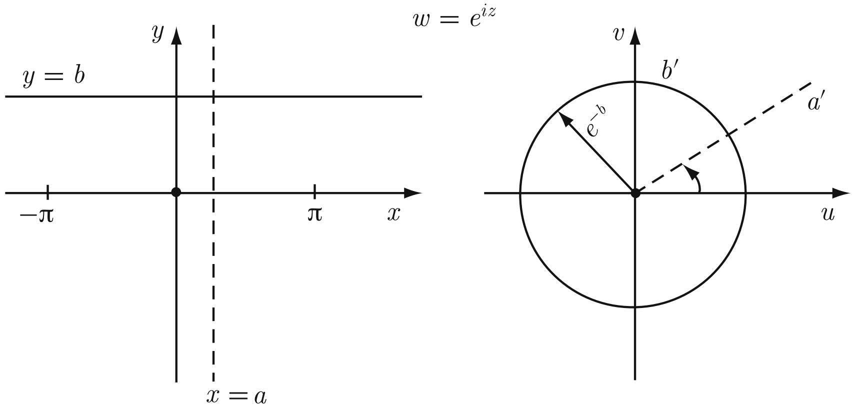


Figure 4.6. Image of the line $x = a$ under e^{iz}

is mapped onto the interior of the punctured unit disk, while the strip

$$\{z : -\pi < x \leq \pi,\, y < 0\},$$

is mapped onto its exterior (see Figure 4.6). Also, the line $x = x_0$, $-\pi < x_0 \leq \pi$, is mapped onto the ray $\operatorname{Arg} w = x_0$.

We will use these mapping properties of the exponential to determine those of the trigonometric functions. We will now discuss the complex mapping $w = \cos z$ of the z-plane onto the w-plane. As was the case with the exponential, we wish to restrict $\cos z$ to a region where the function is one-to-one. Because $\cos z$ is periodic with a real period of 2π, this function assumes all values in any infinite vertical strip $\{z : \alpha < \operatorname{Re} z \leq \alpha + 2\pi\}$. Therefore, it suffices to study the mapping $w = \cos z$ on the strip where α is fixed to be $-\pi$. That is on the strip $\{z : -\pi < \operatorname{Re} z \leq \pi\}$. Note that $\cos(\pi/2) = 0 = \cos(-\pi/2)$ showing that $\cos z$ is not one-to-one on this region. Moreover, $\cos z$ is an even function; that is, $\cos z = \cos(-z)$. This means that the points in the first and fourth quadrants (second and third quadrants) have identical images. Also, it follows that the image of the strip $\{z : -\pi < \operatorname{Re} z < 0\}$ is the same as that of the strip $\{z : 0 < \operatorname{Re} z < \pi\}$, under $w = \cos z$. Hence the image of any set contained in the semi-infinite strip

$$\{z : -\pi < \operatorname{Re} z \leq \pi,\, \operatorname{Im} z > 0\}$$

will be duplicated in any semi-infinite strip of the form

$$\{z : (k-1)\pi < \operatorname{Re} z \leq (k+1)\pi,\, \pm\operatorname{Im} z > 0\}.$$

Since $\cos z$ is real if and only if z is real, it suffices to consider the mapping defined in the region

$$\{z : -\pi < \operatorname{Re} z \leq \pi,\, \operatorname{Im} z > 0\} \cup \{z : 0 \leq \operatorname{Re} z \leq \pi,\, \operatorname{Im} z = 0\},$$

where the function $w = \cos z$ is one-to-one.

Recall that the function $w = (1/2)(z + 1/z)$ maps circles onto ellipses and rays onto arcs of hyperbolas. We may view the transformation

$$w = \cos z = \frac{1}{2}\left(e^{iz} + \frac{1}{e^{iz}}\right)$$

as successive mappings from the z plane to the ζ plane and the ζ plane to the w plane, where

$$\zeta = e^{iz} = e^{-y}e^{ix}.$$

For any $y_0 > 0$, the line segment $-\pi < x \leq \pi$, $y = y_0$, in the z plane is mapped onto the circle $|\zeta| = e^{-y_0}$ in the ζ plane. Then the function

$$w = \frac{1}{2}\left(\zeta + \frac{1}{\zeta}\right) := \frac{1}{2}\left(e^{-y_0}e^{ix} + e^{y_0}e^{-ix}\right)$$

maps the circle $|\zeta| = e^{-y_0}$ in the ζ plane onto the ellipse

$$\left(\frac{u}{\frac{1}{2}(e^{-y_0} + e^{y_0})}\right)^2 + \left(\frac{v}{\frac{1}{2}(e^{-y_0} - e^{y_0})}\right)^2 = 1$$

in the w plane. Hence, $w = \cos z$ maps the line segment $-\pi < x \leq \pi$, $y_0 > 0$ onto an ellipse (see Figure 4.7).

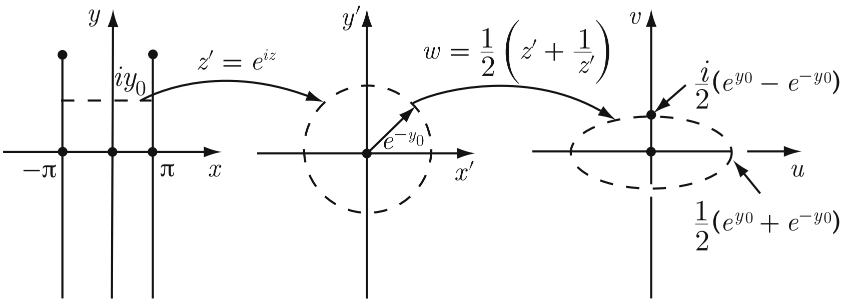


Figure 4.7. Image of the line segment $-\pi < x \leq \pi$, $y_0 > 0$ under $\cos z$

Similarly, for $y > 0$, the half-line $\{z = x_0 + iy : \ y > 0\}$ (where $x_0 \in (-\pi, \pi)$ is fixed), is mapped onto the line segment

$$\operatorname{Arg}\zeta = \operatorname{Arg}(e^{-y}e^{ix_0}) = x_0, \quad 0 < |\zeta| < 1,$$

which, in turn, is mapped onto an arc of the hyperbola (see Figure 4.8)

$$\left(\frac{u}{\cos x_0}\right)^2 - \left(\frac{v}{\sin x_0}\right)^2 = 1.$$

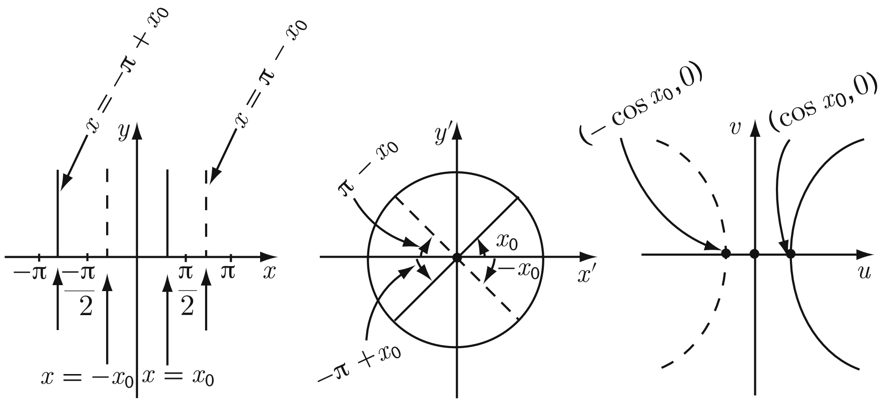


Figure 4.8. Illustration for mapping properties of $\cos z$

Remark 4.9. For $y > 0$, the above mappings are not valid for the half-lines $x = 0$, $x = \pi$, or $x = \pm\pi/2$. If $x = 0$,

$$\cos z = \frac{1}{2}(e^{-y} + e^{y})$$

is mapped onto the real interval $u > 1$, while the half-line $x = \pi$ gives

$$\cos(\pi + iy) = -\frac{1}{2}(e^{-y} + e^{y})$$

so that under $w = \cos z$, $x = \pi$ is mapped onto the real interval $u < -1$. Similarly, for $y > 0$, as

$$\cos\left(\frac{\pi}{2} + iy\right) = \frac{i}{2}(e^{-y} - e^{y})$$

and

$$\cos\left(-\frac{\pi}{2} + iy\right) = -\frac{i}{2}(e^{-y} - e^{y}),$$

the half-line $x = \pi/2$ is mapped onto the negative imaginary axis, and the half-line $x = -\pi/2$ is mapped onto the positive imaginary axis. Finally, as

$$\cos(x + i0) = \frac{e^{ix} + e^{-ix}}{2},$$

the interval $0 \le x \le \pi$, $y = 0$, is mapped onto the real interval $-1 \le u \le 1$, $v = 0$. ●

The identity

$$\sin z = \cos\left(z - \frac{\pi}{2}\right) \tag{4.12}$$

enables us to deduce mapping properties of $\sin z$ from those of $\cos z$. Equation (4.12) shows that we may view the transformation $w = \sin z$ as the translation from the z plane to the ζ plane, where $\zeta = z - \pi/2$, followed by the mapping $w = \cos\zeta$. Thus the function

$$w = \sin z$$

maps points in the region $-\pi/2 < \operatorname{Re} z \le 3\pi/2$ in the same manner as

$$w = \cos z$$

maps the point in the region $-\pi < \operatorname{Re} z \le \pi$.

For instance, the function $w = \cos z$ maps the line segment

$$-\pi < x \le \pi, \quad y = 1,$$

onto the ellipse

$$\frac{u^2}{\frac{1}{4}(e + 1/e)^2} + \frac{v^2}{\frac{1}{4}(e - 1/e)^2} = 1.$$

The function $w = \sin z$ maps the line segment $-\pi/2 < x \le 3\pi/2$, $y = 1$, onto the same ellipse.

Finally, we remark that the relationship between the complex trigonometric and hyperbolic functions, for example,

$$\cos(iz) = \cosh z \quad \text{and} \quad \sin(iz) = i \sinh z,$$

allow us to discuss the action of hyperbolic functions as complex mappings.

Questions 4.10.

1. What kind of function might map the complex plane, excluding the origin, onto the strip $\{w : -\pi < \operatorname{Im} w \le \pi\}$?
2. Given a point in the z plane, does there always exist a neighborhood of that point in which the function e^z is one-to-one?
3. How would you describe the behavior of e^z as z approaches ∞?
4. For the function e^z, how does the area of a rectangle compare with the area of its image?
5. For the function $\cos z$, how does the area of a rectangle compare with the area of its image?
6. For the function e^z, how does the slope of a straight line affect the logarithmic spiral onto which it is mapped?
7. What functions, other than e^z, are never zero in the plane?
8. What is the largest region in which $\sin z$ is one-to-one?
9. Given a point in the z plane, does there always exist a neighborhood of that point in which the function $\sin z$ is one-to-one?
10. What are the differences between the functions $w = \cos(z - \pi/2)$ and $w = \cos z - \pi/2$?

11. What is the image of the infinite strip $\{z : 0 \leq \operatorname{Im} z \leq \pi\}$ under the mapping $w = e^z$?
12. What is the image of the infinite strip $\{z : 0 \leq \operatorname{Im} z \leq \pi/2\}$ under the mapping $w = e^z$?
13. What is the image of the disk $\{z : |z| \leq \pi\}$ under the mapping $w = e^z$?

Exercises 4.11.

1. Find the image of the following sets under the transformation $w = e^z$, and sketch:
   (a) $-5 \leq x \leq 5$, $y = \pi/4$ (b) $x = 3$, $-\pi/2 < y < \pi/2$
   (c) $-2 < x < 1, 0 < y < \pi$ (d) $x < 1$, $-\pi/3 < y < 2\pi/3$.
2. Find the image of the region $0 \leq x \leq \pi$, $y \geq 0$, for the transformation
   (a) $w = e^{iz}$ (b) $w = ie^{iz}$ (c) $w = ie^{-iz}$.
3. Find the images of the straight lines for the transformation $w = e^{cz}$, c a complex constant.
4. Show that the image of the disk $|z| \leq 1$ under the transformation $w = e^z$ is contained in the annulus $1/e \leq |w| \leq e$.
5. Show that the image of the disk $|z| \leq 1$ under the transformations $w = \cos z$ and $w = \sin z$ are contained in the disk $|w| \leq (e^2 + 1)/2e$.
6. Find the image of the following sets under the transformation $w = \cos z$.
   (a) $x = \frac{\pi}{2}$, $y \geq 0$ (b) $-\frac{\pi}{4} < x < \frac{\pi}{4}$, $y = -5$
   (c) $0 \leq x < \pi$, $-2 < y < 2$ (d) $-\frac{\pi}{2} < x < \frac{\pi}{4}$, $y > 0$.

4.3 The Logarithmic Function

Before defining the logarithm of a complex number, we review some properties of the real-valued logarithm. For every positive real number x, there exists a unique real number y such that $e^y = x$. We write $y = \ln x$, and observe that for $x_1, x_2 > 0$, we have

$$\ln(x_1 x_2) = \ln x_1 + \ln x_2.$$

The function $y = \ln x$ maps the positive real numbers onto the set of reals, and is the inverse of the function $y = e^x$, which maps the real numbers onto the positive reals (see Figure 4.9). Since e^x is one-to-one, its inverse is also a one-to-one function.

There is a problem in defining the logarithm of a complex number z as the value w for which

$$e^w = z.$$

We know that for $z = 0$ this equation has no solution in $\mathbb{C}$ because if $w = u+iv$, then one has

$$|e^w| = |e^{u+iv}| = e^u > 0.$$

Thus, e^w never assumes zero in $\mathbb{C}$. Further, the expression

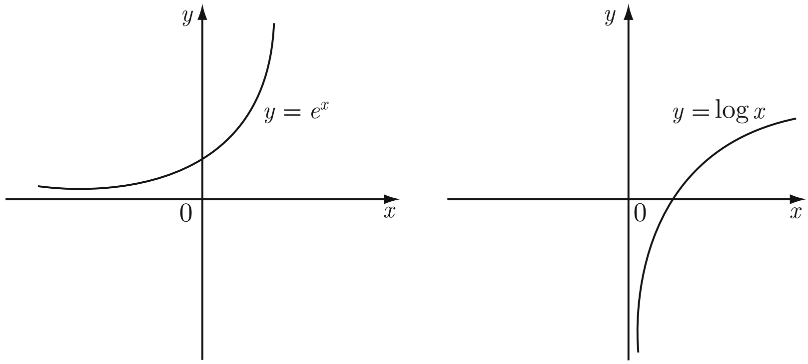


Figure 4.9. Mappings of e^x and $\log x$ for x real

$$e^w = e^u e^{iv}$$

clearly shows that the range of e^w is $\mathbb{C} \setminus \{0\}$. The periodicity of the complex exponential precludes the existence of a *unique* complex logarithm. For, if $e^w = z$, then $e^{w+2k\pi i} = z$ for any integer k. We thus define the logarithm of a complex number z, denoted by $\log z$, as the set of *all* values $w = \log z$ for which $e^w = z$. Thus, as in the real case, for $z \neq 0$

$$w = \log z \Longleftrightarrow e^w = z; \quad \text{or} \quad \log z \in \{w : \, e^w = z\}.$$

Since the exponential function never vanishes, there is no logarithm associated with the complex number zero; and since the exponential assumes every nonzero complex number infinitely often, there are infinitely many values of the logarithm associated with each nonzero complex number. More precisely, we have

Proposition 4.12. *Given $z \neq 0$, the most general solution of $e^w = z$ is given by*

$$w = \log z = \ln |z| + i(\operatorname{Arg} z + 2k\pi i) := \ln |z| + i \arg z, \quad k \in \mathbb{Z}. \tag{4.13}$$

(Remember that there is no solution to $e^w = 0$).

Proof. Setting $z = re^{i\theta}$ $(r > 0, \theta = \operatorname{Arg} z)$, we conclude from

$$e^w := e^u e^{iv} = z = re^{i\theta}$$

that $e^u = r$ and $e^{iv} = e^{i\theta}$, or equivalently

$$u = \ln r \quad \text{and} \quad v = \theta + 2k\pi, \quad k \in \mathbb{Z},$$

where $\ln r$ is the natural logarithm, to the base e, of a positive real number. Therefore, we have the expression

$$w = u + iv = \log z := \ln |z| + i(\operatorname{Arg} z + 2k\pi), \quad k \in \mathbb{Z}$$

which has infinitely many values at each point $z \neq 0$. ■

Remark 4.13. The exponential function e^z has one more important special property. Recall that

$$e^x \to +\infty \quad \text{as } x \to +\infty.$$

Is this true if we replace the real x by a complex z? The answer is clearly no! Indeed, given $0 \neq z \in \mathbb{C}$, there exists a w such that $e^w = z$. But then

$$e^{w+2k\pi i} = z$$

holds for every $k \in \mathbb{Z}$. Hence we can obtain w having arbitrarily large modulus $|w|$ such that $e^w = z$. As a consequence, we conclude that $\lim_{w\to\infty} e^w$ does not exist. Note that

$$\lim_{\substack{w=u\\ u\to\infty}} e^w = \infty, \qquad \lim_{\substack{w=u<0,\\ u\to-\infty}} e^w = 0$$

whereas the limit

$$\lim_{\substack{w=iv\\ v\to\infty}} e^w = \lim_{v\to\infty} e^{iv}$$

does not exist. (For instance, both $v_n = n\pi$ and $v_n' = n\pi + \pi/2$ approach ∞ as $n \to \infty$ but $e^{iv_n} = (-1)^n$ and $e^{iv_n'} = 0$). ●

Since $\log z$ is not a uniquely defined function of z, it is appropriate to introduce the principle value of $\log z$ for $z \neq 0$. For $z \neq 0$,

$$\ln|z| + i\operatorname{Arg} z$$

is called the principle value of $\log z$ and is denoted by $\operatorname{Log} z$:

$$\operatorname{Log} z = \ln|z| + i\operatorname{Arg} z.$$

Using this we can rewrite (4.13) in the form

$$\log z = \operatorname{Log} z + 2k\pi i, \quad k \in \mathbb{Z}.$$

We remark that the expression $w = \log z$, $z \neq 0$, is our first example of a *multiple-valued function*, a relation in which there is more than one image associated with a complex value. Note that the multiple-valuedness of the logarithm is related to the many values connected with the argument of a complex number.

Our methods of investigating continuity and other properties for single-valued functions cannot be used for multiple-valued functions. Fortunately, a multiple-valued function can quite naturally be replaced by many different single-valued functions. The nature of multiple-valued function may then be examined from the point of view of its single-valued counterparts.

We define a *branch* of $\log z$ to be any single-valued function $\log^* z$ that satisfies the identity $e^{\log^* z} = z$ for all nonzero complex values of z. There are infinitely many branches associated with the multiple-valued function $\log z$. Each is an inverse of the function e^z.

Among all the branches for $\log z$, there is exactly one whose imaginary part ($\arg z$) is defined in the interval $(-\pi, \pi]$. This branch is called the *principal branch* of $\log z$ and is $\operatorname{Log} z$. Every branch of $\log z$ differs from the principal branch by a multiple of $2\pi i$. That is, if $\log^* z$ is a fixed branch, then

$$\log^* z = \operatorname{Log} z + 2k\pi i \tag{4.14}$$

for some integer k. Note that

$$\operatorname{Log} z = \ln|z| + i\operatorname{Arg} z \quad (-\pi < \operatorname{Arg} z \leq \pi). \tag{4.15}$$

The restriction in (4.15) may be viewed geometrically as a cut of the z-plane along the negative real axis. This ray is then called the *branch cut* for the function $\operatorname{Log} z$. Other branches of $\log z$ may be defined by restricting $\arg z$ to

$$(2k-1)\pi < \arg z \leq (2k+1)\pi, \quad k \text{ an integer.}$$

The "cut line" may not be crossed while continuously varying the argument of z without moving from one branch to another, which would destroy single-valuedness.

Undue importance should not be placed on the restrictions of the argument in (4.15) to the interval $(-\pi, \pi]$. For a fixed α real, the function

$$\log_\alpha z = \ln|z| + i\arg z \quad (\alpha < \arg z \leq \alpha + 2\pi),$$

which has branch cut $\arg z = \theta + 2\pi$, would serve our purposes just as well. Indeed, a branch cut need not even be confined to a ray. Any continuous nonself-intersecting curve that extends from the origin to infinity would do.

We next examine the extent to which the identity

$$\log(z_1 z_2) = \log z_1 + \log z_2 \quad (z_1, z_2 \neq 0)$$

is valid in the complex plane. For a fixed branch, we have

$$\log(z_1 z_2) = \ln|z_1 z_2| + i\arg(z_1 z_2)$$

and

$$\begin{aligned}\log z_1 + \log z_2 &= \ln|z_1| + \ln|z_2| + i(\arg z_1 + \arg z_2)\\ &= \ln|z_1 z_2| + i(\arg z_1 + \arg z_2).\end{aligned}$$

As we have seen in Section 1.3, $\arg(z_1 z_2)$ differs from $\arg z_1 + \arg z_2$ by an integer multiple of 2π. Thus, the best we can do is

$$\log(z_1 z_2) = \log z_1 + \log z_2 + 2k\pi i$$

or

$$\log(z_1 z_2) = \log z_1 + \log z_2 \pmod{2\pi i}.$$

When the cut is along the negative real axis, the branch of the logarithm under consideration is completely determined by specifying one particular value of the function. For instance, the principal branch is the only one for which $\log 1 = 0$. The branch for which $\log 1 = 10\pi i$ is given by

$$\log z = \operatorname{Log} z + 10\pi i.$$

Each of the functions $w = \operatorname{Log} z + 2k\pi i$ maps the plane, excluding the origin, onto the infinite strip $(2k-1)\pi < \operatorname{Im} w \leq (2k+1)\pi$. Recall that the exponential function maps each strip

$$(2k-1)\pi < \operatorname{Im} w \leq (2k+1)\pi$$

onto the punctured plane. Since the behavior of each function defined in (4.14) is essentially the same, we will—unless otherwise stated—assume $k = 0$ and confine ourselves to the principal branch of the logarithm.

The function $w = \operatorname{Log} z$ is not continuous at any point on the negative real axis. For any such point may be expressed as

$$z_0 = r_0 e^{\pi i}, \; r_0 > 0,$$

with $\operatorname{Log} z_0 = \ln r_0 + i\pi$. But as the point z_0 is approached through values below the real axis, we have

$$\lim_{z \to z_0} \operatorname{Arg} z = -\pi.$$

Hence,

$$\operatorname{Log} z \to \ln r_0 - i\pi \neq \operatorname{Log} z_0 \quad \text{as} \quad z \to z_0$$

through such values.

This does not mean that the logarithm function is not continuous on the negative real axis. All we have seen is that $\operatorname{Log} z$, the principal branch, is not continuous at these points. By making our cut along a different ray, we can find a branch of the logarithm that is continuous for negative real values. For instance, the single-valued function

$$w = \log z = \ln |z| + i \arg z \quad (-\pi/2 < \arg z \leq 3\pi/2)$$

is continuous at all points on the negative real axis, but not on the ray $\arg z = 3\pi/2$.

In other words, the logarithm function is continuous for all nonzero complex values in the following sense: Given $z_0 \neq 0$, there exists a branch for which $\lim_{z \to z_0} \log z = \log z_0$. However, there does not exist a branch for which $\log z$ is continuous for all nonzero complex numbers.

In view of (4.15), we can easily determine some mapping properties of the logarithm function. The image of the circle $|z| = r$ for the function $w = \operatorname{Log} z$ is the line segment $u = \ln r$, $-\pi < v \leq \pi$ (see Figure 4.10). We also have the ray $\operatorname{Arg} z = \theta$ mapping onto the line $v = \theta$ (see Figure 4.11).

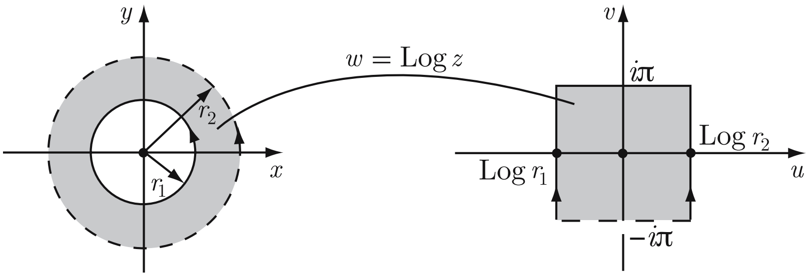


Figure 4.10. Image of an annulus region under $\operatorname{Log} z$

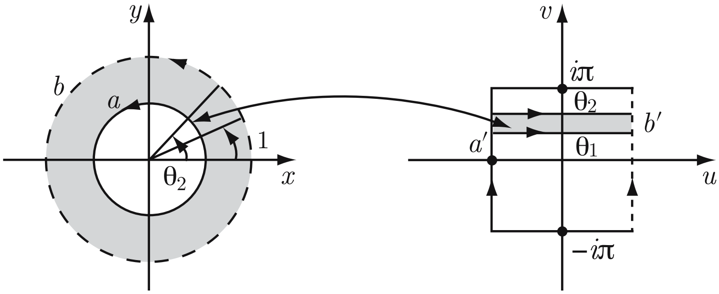


Figure 4.11. Image of segment of rays under $\operatorname{Log} z$

Questions 4.14.

1. What is the relationship between the argument and the logarithm of a complex number?
2. For fixed θ_0, what changes will occur if we define $\theta_0 - \pi < \arg z \leq \theta_0 + \pi$ to be the principal value?
3. What would be the consequences of defining $\log 0 = \infty$?
4. What is the image of spirals under the function $w = \log z$?
5. In what regions is $\log z$ bounded?
6. Does $\log(z_1/z_2) = \log z_1 - \log z_2$?
7. Does $\operatorname{Log}(z_1/z_2) = \operatorname{Log} z_1 - \operatorname{Log} z_2$?
8. Does $\lim_{z\to\infty} \exp(-z^2)$ exist? Does $\lim_{z\to\infty} \exp(-z^4)$ exist?
9. Does $\lim_{z\to 0} \exp(-1/z)$ exist? Does $\lim_{z\to 0} \exp(-1/z^2)$ exist? How about $\lim_{z\to 0} \exp(-1/z^3)$? Does $\lim_{z\to 0} \exp(-1/z^4)$ exist?

Exercises 4.15.

1. Find all the values of
   (a) $\log(1-i)$ (b) $\log(3-2i)$ (c) $\log(x+iy)$.
2. For any nonzero complex number z_1 and z_2, prove that

$$\operatorname{Log}(z_1 z_2) = \operatorname{Log} z_1 + \operatorname{Log} z_2 + 2k\pi i,$$

where $k = 0, 1, -1$. Give examples to show that each value of k is possible.

3. For $z \neq 0$, prove that $\ln|z| \leq |\operatorname{Log} z| \leq \ln|z| + |\operatorname{Arg} z|$.
4. Let $f(z)$ be defined in a domain D with $f(z) \neq 0$ in D. Prove that $\operatorname{Arg} f(z) = \operatorname{Im} \operatorname{Log} f(z)$ for every point $z \in D$.
5. Find the image of straight lines parallel to the coordinate axes for the function
   (a) $w = \operatorname{Log}(iz)$ (b) $w = \operatorname{Log}(-iz) + 1$ (c) $w = \operatorname{Log} z^2$.
6. Set $z = re^{i\theta}$, where $0 < a < r < b$ and $\theta \in (-\pi, \pi)$. Show that under the mapping $\operatorname{Log} z$ the region $\Omega = \{z : a < |z| < b\} \setminus [-b, -a]$ is mapped onto the rectangle $(\ln a, \ln b) \times (-\pi, \pi)$.

4.4 Complex Exponents

As we have seen in Section 1.3 there are n distinct complex values associated with $z^{1/n}$, $z \neq 0$. If we write $z = re^{i\theta}$, then

$$z_k = r^{1/n} e^{i[(\theta+2k\pi)/n]}$$

is a distinct nth root for $k = 0, 1, 2, \ldots, n-1$. The values of $z^{1/n}$ vary as the argument of z takes on the values $\theta, \theta + 2\pi, \theta + 4\pi, \ldots, \theta + 2(n-1)\pi$. A unique value for $z^{1/n}$ may be obtained by restricting $\arg z$ to a particular branch. This seems to indicate a link between the nth roots and logarithm of a nonzero complex number.

Indeed, we may define the function $z^{1/n}$ by

$$w = z^{1/n} = e^{(1/n)\log z} = e^{(1/n)(\ln|z| + i\arg z)}. \tag{4.16}$$

This function, like the logarithm function, is multiple-valued. Upon setting $\arg z = \operatorname{Arg} z + 2k\pi$, k an integer, we see that (4.16) assumes different values for $k = 0, 1, \ldots, n-1$; for any other integer k, one of these n values will be repeated. More generally, if m and n are positive integers with no common factors, we define

$$(z^{1/n})^m = e^{(m/n)\log z}, \quad z \neq 0.$$

This, too, has n distinct values. We are thus led quite naturally to define z^α for complex values of α by

$$z^\alpha = e^{\alpha \log z}. \tag{4.17}$$

If α is not a rational number, then there are infinitely many values associated with the expression in (4.17). To see this, we first suppose that α is irrational. Then for $z = re^{i\theta} \neq 0$, we have

$$z^\alpha = e^{\alpha \log z} = e^{\alpha[\ln r + i(\theta+2k\pi)]} = r^\alpha e^{i\alpha\theta} e^{i(2k\pi\alpha)}.$$

Since

$$e^{i(2k_1\pi\alpha)} = e^{i(2k_2\pi\alpha)} \Longleftrightarrow \alpha(k_1 - k_2) = m \ \ (m \in \mathbb{Z}),$$

we see that α is real rational number which is a contradiction. It follows that z^α has infinitely many distinct values, all with the same modulus.

Next suppose that $\alpha = a + ib$ (a and b real, $b \neq 0$). Then

$$z^{a+ib} = e^{(a+ib)\log z} = e^{a\ln r - b(\theta+2k\pi)} e^{i(b\ln r + a\theta + 2k\pi a)}.$$

Since $|z^{a+ib}| = r^a e^{-b(\theta+2k\pi)}$, the complex number z^{a+ib} has a different modulus for each branch, any two of which differ by a factor of $e^{-2n\pi}$, n an integer.

Examples 4.16. We have

(i) $5^{1/2} = e^{(1/2)\log 5} = e^{(1/2)(\ln 5 + 2k\pi i)} = e^{(1/2)\ln 5} e^{k\pi i} = \pm\sqrt{5}$

(ii) $i^{1/2} = e^{(1/2)\log i} = e^{(1/2)i(\pi/2+2k\pi)} = \pm e^{\pi i/4} = \pm\frac{\sqrt{2}}{2}(1+i)$

(iii) $i^i = e^{i\log i} = e^{i[\ln 1 + i(\pi/2+2k\pi)]} = e^{-(\pi/2+2k\pi)}$,

where k is any integer. ●

Example 4.17. To find all possible solutions of $z^{1-i} = 4$, we rewrite $z^{1-i} = 4$ as

$$e^{(1-i)\log z} = 4 = e^{\ln 4 + 2k\pi i} \quad (k \in \mathbb{Z})$$

so that

$$(1-i)\log z = 2\ln 2 + 2k\pi i, \quad \text{i.e.,} \quad \log z = [\ln 2 - k\pi] + i[\ln 2 + k\pi].$$

By the definition of $\log z$, we have

$$z = e^{[\ln 2 - k\pi] + i[\ln 2 + k\pi]}, \quad k \in \mathbb{Z}.$$

Simplification of this relation shows that the solutions of $z^{1-i} = 4$ are given by $z = 2e^{-k\pi}e^{i[\ln 2 + k\pi]}$, $k \in \mathbb{Z}$. ●

Remark 4.18. In some contexts, the expression $x^{1/2}$ and $\sqrt{x}$ are used interchangeably. By $x^{1/2} (x > 0)$, we mean both the positive real number $+\sqrt{x}$ and the negative real number $-\sqrt{x}$. ●

It is of interest to compare the relationship between $z^\alpha z^\beta$ and $z^{\alpha+\beta}$, where $z = re^{i\theta}$ $(r \neq 0)$. If α and β are real, then

$$z^\alpha z^\beta = e^{\alpha\log z} e^{\beta\log z} = e^{(\alpha+\beta)\ln r} e^{i(\alpha+\beta)\theta} e^{2\pi i(k\alpha + n\beta)},$$

where k and n are integers. On the other hand,

$$z^{\alpha+\beta} = e^{(\alpha+\beta)\log z} = e^{(\alpha+\beta)\ln r} e^{i(\alpha+\beta)\theta} e^{2\pi i m(\alpha+\beta)}$$

for m an integer. Thus, if α and β are integers, $z^\alpha z^\beta = z^{\alpha+\beta}$. If either α or β is an integer, then $z^\alpha z^\beta$ and $z^{\alpha+\beta}$ assume the same set of values, although equality for each α and β need not hold. In general, $z^{\alpha+\beta}$ assumes every value of $z^\alpha z^\beta$, but the converse is not true. For example $5^{1/2+1/2} = 5$ but $5^{1/2}5^{1/2} = \pm 5$. We leave it for the reader to show this containment for α and β complex numbers.

Recall the "proof" in the introduction that $1 = -1$, where we made the false assumption that

$$\sqrt{1/-1} = \sqrt{1}/\sqrt{-1}.$$

Had we used the preceding results, we would have been able to reach only the much less interesting conclusion that $\pm 1 = \pm 1$.

As was the case with the logarithmic function, we may replace multiple-valued functions that have fractional exponents with (single-valued) branches. To illustrate, for $z = re^{i\theta}$ ($r \neq 0, -\pi < \theta \leq \pi$) the principal branch of $z^{1/2}$ is

$$w_0 = z^{1/2} = e^{(1/2)(\ln r + i\theta)} = \sqrt{r}e^{i(\theta/2)} \quad (-\pi < \theta \leq \pi).$$

Another branch of the function is

$$w_1 = z^{1/2} = e^{(1/2)(\ln r + i(\theta+2\pi))} = -\sqrt{r}e^{i(\theta/2)} \quad (\pi < \theta + 2\pi \leq 3\pi).$$

Both w_0 and w_1 are continuous functions, except on the negative real axis. This ray is called a *branch cut* for both w_0 and w_1. Each of these single-valued functions is called a *determination* or branch of the multiple-valued function $w = z^{1/2}$.

We now establish some mapping properties for the functions w_0 and w_1. The punctured plane ($z \neq 0$) is mapped by w_0 onto the right half-plane, including the positive imaginary axis, and by w_1 onto the left half-plane, including the negative imaginary axis. These functions also map circles onto semicircles, excluding the end point (see Figure 4.12).

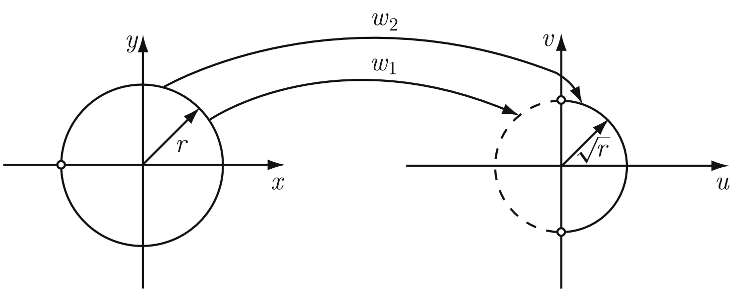


Figure 4.12. Mapping properties of square root function

We may similarly analyze other multiple-valued functions with rational exponents. The function $w = z^{1/3}$ has three branches. We write

$$w_k = \sqrt[3]{r}e^{i(\theta+2k\pi)/3} \quad (k = 0, 1, 2; -\pi < \theta \leq \pi).$$

Each of these three single-valued functions is continuous except on the negative real axis, and maps the circle $|z| = r$ onto the arc

$$|w_k| = \sqrt[3]{r}, \quad \frac{(2k-1)\pi}{3} < \operatorname{Arg} w_k \le \frac{(2k+1)\pi}{3}.$$

The next example illustrates some of the surprising properties of complex exponents. Consider the function $w = 1^z$. We have

$$w = 1^z = e^{z \log 1} = e^{2k\pi i z} = e^{2k\pi i(x+iy)} = e^{-2k\pi y} e^{2k\pi i x}.$$

For each integer k, the function $w = 1^z$ is defined in the whole plane. If $k = 0$, the principal branch of the logarithm, then $w \equiv 1$. This is what we expect.

But consider a different determination of the logarithm, and assume that $k = k_0$, $k_0 > 0$. The function $w = 1^z$ is then periodic, with period $1/k_0$. If z is a positive integer, then $1^z = 1$. If z is real, then 1^z is a point on the unit circle. In fact, every interval of the form

$$x_0 - \frac{1}{2k_0} < x \le x_0 + \frac{1}{2k_0}, \quad x_0 \text{ fixed},$$

maps one-to-one onto the unit circle. The line segment

$$-\frac{1}{2k_0} < x \le \frac{1}{2k_0}, \; y = y_0,$$

maps onto the circle $|w| = e^{-2k_0 \pi y_0}$. The line

$$x = x_0, \quad -\frac{1}{2k_0} < x_0 \le \frac{1}{2k_0},$$

maps onto the ray $\operatorname{Arg} w = 2k_0 \pi x_0$. Hence the infinite strip

$$\left\{ z : -\frac{1}{2k_0} < \operatorname{Re} z \le \frac{1}{2k_0} \right\}$$

maps onto the plane, excluding the origin. Finally, the upper half-plane is mapped onto the interior of the punctured unit disk, and lower half-plane onto its exterior (see Figure 4.13).

We have previously examined the close relation between the exponential and trigonometric functions. It is not surprising that their inverses also have much in common. We will show that the inverse trigonometric functions may be defined in terms of logarithm.

Given a complex number z, we wish to find all the complex numbers w such that $z = \sin w$. If w_0 is one such solution, then $w_0 + 2k\pi$ must also be a solution. If

$$z = \sin w = (e^{iw} - e^{-iw})/2,$$

then

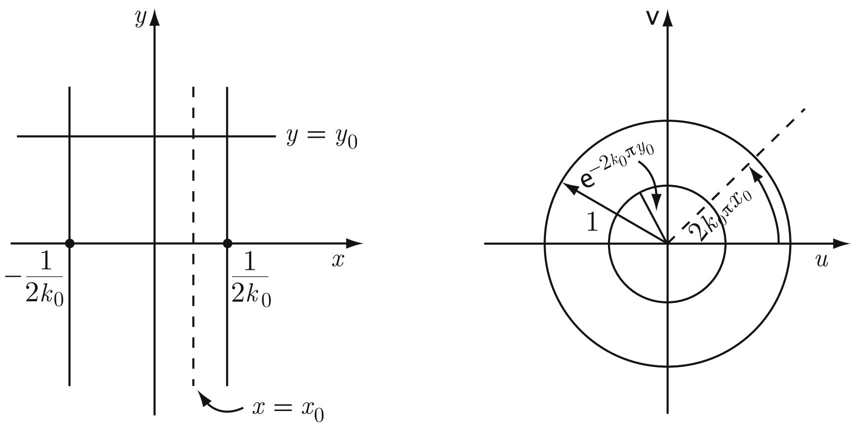


Figure 4.13. Illustration for mapping properties of 1^z

$$e^{2iw} - 2ize^{iw} - 1 = 0,$$

a quadratic in e^{iw}. Solving, we obtain

$$e^{iw} = iz + (1 - z^2)^{1/2}; \quad \text{i.e., } iw = \log[iz + (1 - z^2)^{1/2}],$$

from which we define the multiple-valued function

$$w = \sin^{-1} z = -i \log[iz + (1 - z^2)^{1/2}].$$

Remark 4.19. The expression $(1 - z^2)^{1/2}$ is itself multiple-valued. To compute $\sin^{-1} z$, we must first find each of the values of $(1 - z^2)^{1/2}$. For each of these values, we must then determine all the logarithms. ●

Example 4.20. To find all possible determinations of $\sin^{-1} 0$, we write

$$\sin^{-1} 0 = -i \log 1^{1/2} = -i\frac{2k\pi i}{2} = k\pi, \quad k \in \mathbb{Z}.$$

For the principal branch, $k = 0$ and $\sin^{-1} 0 = 0$. ●

Example 4.21. To find all possible determinations of $\sin^{-1} i$, we write

$$\begin{aligned}\sin^{-1} i &= -i \log(-1 + 2^{1/2})\\ &= -i\left[\ln|-1 \pm \sqrt{2}| + i \arg(-1 \pm \sqrt{2})\right]\\ &= \arg(-1 \pm \sqrt{2}) - i \ln|-1 \pm \sqrt{2}|.\end{aligned}$$

Given any determination for the square root and logarithm, the real part of $\sin^{-1} i$ is an integral multiple of π. The imaginary part depend only on the determination of the square root. Hence

$$\sin^{-1} i = k\pi - i \ln|-1 \pm \sqrt{2}|, \quad k \in \mathbb{Z}.$$

A choice of the positive square root and the principal value for the logarithm gives the specific determination $\sin^{-1} i = -i \ln(\sqrt{2} - 1)$.

We may similarly find the inverses of the remaining trigonometric functions. For $z = \cos w = (e^{iw} + e^{-iw})/2$, we have

$$w = \cos^{-1} z = -i \log[z + (z^2 - 1)^{1/2}].$$ ●

Questions 4.22.

1. Does $\log z^\alpha = \alpha \log z$ if $\alpha = 0$?
2. Does $1^{1/2} + 1^{1/2} = 2(1^{1/2})$?
3. When will a complex number to a complex power be real?
4. If z^α assumes m distinct values and z^β assumes n distinct values, what can we say about $z^\alpha z^\beta$?
5. For the complex number α and β, how does $z^{\alpha\beta}$ compare with $(z^\alpha)^\beta$?
6. For the multiple-valued function $w = z^{1/2}$, why does the origin play such an important role?
7. For the function $w = z^{1/2}$, could we have chosen any rays other than $\operatorname{Arg} z = \pi$ for our branch cut?
8. How do the functions $w = z^{1/n}$ and $w = z^n$ compare?
9. Does $(z^m)^{1/n} = (z^{1/n})^m$ when m and n are integers?
10. When is $\sqrt{z^2} = z$? Is $(z^2)^{1/2} = z$?
11. In what regions are the inverse trigonometric functions one-to-one?
12. How can mapping properties of the inverse trigonometric functions be determined from those of the trigonometric functions?
13. If $|\cos z| \le 1$, then what can you say about z?
14. If $\cos z = \alpha$, where $-1 \le \alpha \le 1$, then what can you say about z?
15. If $\log 4$ is real, what must be the value of $\log(4i)$? What must be the value of $\log(-4i)$?

Exercises 4.23.

1. Find all values for the following expressions.
   (a) 5^i (b) $(\pi i)^e$ (c) $(2^i)^i$ (d) $\log(1+i)^{\pi i}$.
2. For $z \ne 0$, α and β complex numbers, show that every value of $z^{\alpha\beta}$ is a value of $(z^\alpha)^\beta$. When is the converse true?
3. For $z \ne 0$ and α irrational, show that $\theta_0 < \operatorname{Arg}(z^\alpha) < \theta_0 + \epsilon$ for infinitely many values of z^α, where $-\pi < \theta_0 < \pi$ and $\epsilon > 0$.
4. Separate into real and imaginary parts.
   (a) x^x (x real, $x \ne 0$) (b) $(iy)^{iy}$ (y real, $y \ne 0$) (c) z^z ($z \ne 0$).
5. For any nonzero complex number a, show that a^z is either constant or an unbounded function, depending on the branch chosen for its logarithm.

6. Discuss the image of the circle $|z| = r$ for the following multiple-valued functions.
   (a) $w = z^{1/n}$, n a positive integer (b) $w = z^{2/3}$.
7. Prove the following identities:
   (a) $\tan^{-1} z = \dfrac{1}{2i} \log \left(\dfrac{1+zi}{1-zi} \right)$
   (b) $\cot^{-1} z = \dfrac{1}{2i} \log \left(\dfrac{z+i}{z-i} \right)$
   (c) $\sec^{-1} z = \dfrac{1}{i} \log \left(\dfrac{1+(1-z^2)^{1/2}}{z} \right)$
   (d) $\csc^{-1} z = \dfrac{1}{i} \log \left(\dfrac{i+(z^2-1)^{1/2}}{z} \right)$.
8. Find all values of

$$\text{(a)} \quad \sin^{-1} \frac{1}{2} \quad \text{(b)} \quad \cos^{-1} \frac{\sqrt{2}}{2} \quad \text{(c)} \quad \tan^{-1}(1+i) \quad \text{(d)} \quad \sec^{-1} i.$$

9. Determine all values of the following:

$$\begin{array}{lllll} (-1)^{\sqrt{2}}, & 2^{1-i}, & (1+i)^{\sqrt{3}}, & \arg(1-i), & (-1)^{1/3} \\ (\cos i)^i, & (1+i)^{1+i}, & (i^i)^i, & i^{\sin i}, & (\sqrt{3}+i)^{i/2}. \end{array}$$

10. Evaluate the limits

$$\text{(i)} \lim_{z \to 0} (1+z)^{1/z} \qquad \text{(ii)} \lim_{z \to 2} \frac{\sqrt{z}-2}{z-2} \qquad \text{(iii)} \lim_{z \to i} z \operatorname{Arg}(\overline{z}).$$

11. Find all the points of discontinuity of

    (i) $f(z) = \operatorname{Log}(z^2-1)$ (ii) $f(z) = \operatorname{Arg}(z^2)$
    (iii) $f(z) = \operatorname{Log}(z^3-1)$ (iv) $f(z) = \operatorname{Arg}(z^3)$
    (v) $f(z) = \sqrt{z^2+1}$ (vi) $f(z) = \sqrt{z^2-1}$.

5

Analytic Functions

In this chapter, we will define differentiation for single-valued functions of a complex variable, and we will see how the derivative of a complex variable sometimes behaves like the derivative of a real function of one real variable, and other times it is comparable to the partial derivatives of a real function of two real variables. We also learn to appreciate the importance of neighborhoods. If we do not require differentiability in a neighborhood, the smoothness of a function along one path may obscure potential difficulties along some other route.

5.1 Cauchy–Riemann Equation

As we have seen before, a function of a complex variable may be separated into its real and imaginary parts. Writing $f(z) = u(x, y) + iv(x, y)$ it is interesting to compare properties of $f(z)$ with those of its real-valued components $u(x, y)$ and $v(x, y)$. In the case of continuity, the comparison is quite straight forward.

Let g be a function of the two real variables x and y. We say that

$$\lim_{(x,y)\to(x_0,y_0)} g(x, y) = L$$

if for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|g(x, y) - L| < \epsilon \quad \text{whenever } 0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta.$$

If $L = g(x_0, y_0)$, then $g(x, y)$ is said to be *continuous* at (x_0, y_0).

Theorem 5.1. *The function $f(z) = u(x, y)+iv(x, y)$ is continuous at a point $z_0 = x_0+iy_0$ if and only if $u(x, y)$ and $v(x, y)$ are both continuous at the point (x_0, y_0).*

Proof. We first suppose $f(z)$ to be continuous at $z = z_0$. Then for any $\epsilon > 0$,

$$|u(x,y) - u(x_0,y_0)| \le |f(z) - f(z_0)| < \epsilon,$$
$$|v(x,y) - v(x_0,y_0)| \le |f(z) - f(z_0)| < \epsilon,$$

whenever

$$|z - z_0| = \sqrt{(x-x_0)^2 + (y-y_0)^2} < \delta.$$

Thus both $u(x,y)$ and $v(x,y)$ are continuous at (x_0,y_0). Conversely, if $u(x,y)$ and $v(x,y)$ are continuous at (x_0,y_0), the continuity of $f(z)$ follows from the inequality (recall that $|z| \le |x| + |y|$ for $z \in \mathbb{C}$)

$$|f(z) - f(z_0)| \le |u(x,y) - u(x_0,y_0)| + |v(x,y) - v(x_0,y_0)|. \qquad \blacksquare$$

In view of Theorem 5.1, it is worthwhile to examine more carefully the notion of the limit of a function of real two variables. Recall that for a function of one real variable, when there are only two directions to travel, a limit exists and only if the right- and left-hand limits coincide. There is no analog for a function of two variables, since infinitely many modes of approach are possible.

Example 5.2. Let $f(x,y) = xy/(x^2+y^2)$, $(x,y) \ne (0,0)$. Since $f(x,y) \equiv 0$ as $(x,y) \to (0,0)$ along either of the coordinate axes, we have

$$\lim_{y\to 0} f(0,y) = \lim_{x \to 0} f(x,0) = 0.$$

However, choosing the straight-line path $y = mx$, we obtain

$$\lim_{x\to 0} f(x,mx) = \lim_{x\to 0} \frac{mx^2}{x^2+mx^2} = \frac{m}{1+m^2}.$$

Because $f(x,y)$ approaches different values along different straight lines, the limit at the origin does not exist. $\bullet$

Example 5.3. Let $f(x,y) = x^2y^2/(x+y^2)^3$, $(x,y) \ne (0,0)$. Here,

$$\lim_{x\to 0} f(x,mx) = \lim_{x\to 0} \frac{m^2x^4}{(x+m^2x^2)^3} = \lim_{x\to 0} \frac{m^2x}{(1+m^2x)^3} = 0,$$

and $f(x,y)$ approaches 0as $(x,y) \to (0,0)$ along any straight line. But along the parabola $x = my^2$ $(m \ne 0)$,

$$\lim_{y\to 0} f(my^2,y) = \lim_{y\to 0} \frac{m^2y^4y^2}{(my^2+y^2)^3} = \frac{m^2}{(m^2+1)^3}.$$

Hence, $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist. $\bullet$

Example 5.4. Let $f(x,y) = (x^3 - 2y^3)/(x^2+y^2)$, $(x,y) \ne (0,0)$. We wish to show that this function does have a limit at origin. Obviously, we can not try all modes of approach; so it will be necessary to obtain appropriate inequalities. We have

$$\begin{aligned}|x^3 - 2y^3| &\le |x|^3 + 2|y|^3 = |x|x^2 + 2|y|y^2 \\ &\le \sqrt{x^2+y^2}(x^2+2y^2) \le 2(x^2+y^2)^{3/2}.\end{aligned}$$

Thus, $|f(x,y)| \le 2\sqrt{x^2+y^2}$, so that

$$|f(x,y)| < \epsilon \quad \text{whenever } \sqrt{x^2+y^2} < \epsilon/2 = \delta.$$

Therefore,

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0.$$ ●

Our purpose in the above examples was not to develop sophisticated methods for evaluating limits, but to emphasize the importance of the path. A function may be very well-behaved along one route, while impossible to deal with along another. This phenomenon will produce some surprising results within the theory of differentiable functions.

A function f is said to be *differentiable* at a point z if

$$\lim_{h\to 0} \frac{f(z+h)-f(z)}{h}$$

exists. We then write

$$f'(z) = \lim_{h\to 0} \frac{f(z+h)-f(z)}{h}.$$

Note that h approaches 0 through points in the *plane*, not just along the real axis or along the line $y = mx$. For example, the function $f(z) = z^2$ is everywhere differentiable because

$$\lim_{h\to 0} \frac{f(z+h)-f(z)}{h} = \lim_{h\to 0} \frac{(z+h)^2 - z^2}{h} = \lim_{h\to 0} \frac{2zh + h^2}{h} = 2z.$$

More generally, if $f(z) = z^n$, n an integer, then $f'(z) = nz^{n-1}$. From the identity

$$f(z+h) - f(z) = \left(\frac{f(z+h)-f(z)}{h}\right) h \quad (h \ne 0),$$

we let $h \to 0$ to obtain the familiar real-variable theorem that a differentiable function is continuous. That is,

$$\lim_{h\to 0}(f(z+h) - f(z)) = \lim_{h\to 0} \left(\frac{f(z+h)-f(z)}{h}\right) h = f'(z)\cdot 0 = 0.$$

Therefore, $\lim_{h\to 0} f(z+h) = f(z)$ and so we have

Theorem 5.5. *If f is differentiable at a point $z \in \mathbb{C}$, then f is continuous at z.*

On the other hand, $f(z) = |z| = \sqrt{x^2 + y^2}$ is continuous on $\mathbb{C}$ but not differentiable at the origin. Clearly, $\lim_{z\to 0} f(z) = f(0) = 0$ and

$$\frac{f(h) - f(0)}{h - 0} = \frac{|h|}{h} = \begin{cases} 1 & \text{for positive real values of } h \\ -1 & \text{for negative real values of } h \\ -i & \text{for values of } h \text{ on the positive} \\ & \quad \text{imaginary axis} \\ i & \text{for values of } h \text{ on the negative} \\ & \quad \text{imaginary axis.} \end{cases}$$

Hence, $f'(0)$ does not exist. This example shows that continuity at a point is not sufficient for a function f to be differentiable at that point. Also it follows that each of the functions $\overline{z}$, $\mathrm{Re}\, z$ and $\mathrm{Im}\, z$ is continuous on $\mathbb{C}$ but nowhere differentiable.

Note the similarity thus far between the derivative of the complex-valued function $f(z)$ and the real-valued function $f(x)$. In fact, all the formal differentiation rules are the same. We collect the following properties as

Theorem 5.6. *For $f(z)$ and $g(z)$ differentiable,*

(i) $(f(z) + g(z))' = f'(z) + g'(z)$,
(ii) $(f(z)g(z))' = f'(z)g(z) + f(z)g'(z)$,
(iii) $\left(\dfrac{f(z)}{g(z)}\right)' = \dfrac{f'(z)g(z) - g'(z)f(z)}{(g(z))^2}$ *when* $g(z) \neq 0$.

Suppose g is differentiable at z, and f is differentiable at $g(z)$. If $F(z) = f(g(z))$, then

(iv) $F'(z) = f'(g(z))g'(z)$.

Proof. These properties are proved, word for word, as they are for a function of a real variable. ■

The function $f(z) = |z|^2$ is differentiable only at the origin. In contrast, $g(x) = |x|^2$ $(x \in \mathbb{R})$ is differentiable everywhere in $\mathbb{R}$. Similarly, the function $\phi(x) = x$ is differentiable on $\mathbb{R}$, but the function $\psi(z) = x = \mathrm{Re}\, z$ $(z \in \mathbb{C})$ is nowhere differentiable.

Lest the reader be fooled by the apparent similarity between real and complex derivatives, we now discuss some of the far-reaching consequences of requiring $(f(z + h) - f(z))/h$ to approach the same value regardless of the path of h. To this end, properties of the real and imaginary parts of the differentiable function $f(z) = u(x, y) + iv(x, y)$ will be deduced by specializing the mode of approach.

Suppose first that h approaches 0 along the real axis. Separating into real and imaginary parts,

$$\begin{aligned} \frac{f(z+h) - f(z)}{h} &= \frac{u(x+h, y) + iv(x+h, y) - u(x, y) - iv(x, y)}{h} \\ &= \frac{u(x+h, y) - u(x, y)}{h} + i\frac{v(x+h, y) - v(x, y)}{h}. \end{aligned}$$

If f is differentiable at $z = x + iy$, the limits of both expressions on the right must exist, and we recognize them to be the partial derivatives of $u(x, y)$ and $v(x, y)$ with respect to x. Hence

$$f'(z) = \lim_{h\to 0} \frac{f(z+h) - f(z)}{h} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} := \frac{\partial f}{\partial x}. \tag{5.1}$$

Next let h approach 0 along the imaginary axis. Then for $h = ik$, k real, we have

$$\frac{f(z+ik) - f(z)}{ik} = \frac{u(x, y+k) - u(x, y)}{ik} + i\frac{v(x, y+k) - v(x, y)}{ik}.$$

Thus

$$f'(z) = \lim_{k\to 0} \frac{f(z+ik) - f(z)}{ik} = \frac{1}{i}\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y} = -i\frac{\partial f}{\partial y}. \tag{5.2}$$

But the expressions in (5.1) and (5.2) must be equal. So

$$f'(z) = u_x + iv_x = v_y - iu_y. \tag{5.3}$$

Equating real and imaginary parts in (5.3), we obtain

$$u_x = v_y, \quad v_x = -u_y; \text{ or } f_x = -if_y. \tag{5.4}$$

The two equations in (5.4), known as the *Cauchy–Riemann equations*, furnish us with a necessary condition for differentiability at a point.

Theorem 5.7. *If $f = u + iv$ is differentiable at z, then f_x, f_y exist at z and satisfy the Cauchy–Riemann equations at z:*

$$f_y(z) = if_x(z);$$

or equivalently, $u_x = v_y$ and $u_y = -v_x$.

For the function $f(z) = \overline{z} = x - iy$, we have $u_x \equiv 1$ and $v_y \equiv -1$. Since the Cauchy–Riemann equations are never satisfied, the function is nowhere differentiable.

To see that the Cauchy–Riemann equations are not a sufficient condition for differentiability at a point, consider

$$f(x+iy) = u + iv = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2} + i\left(\dfrac{x^3 + y^3}{x^2 + y^2}\right) & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } x = y = 0. \end{cases}$$

For this function the corresponding u and v are continuous at the origin and the partial derivatives of u and v all exist at the origin, because

$$u_x(0, 0) = \lim_{h\to 0} \frac{u(h, 0) - u(0, 0)}{h} = \lim_{h\to 0} \frac{h^3/h^2 - 0}{h} = 1,$$

$$u_y(0,0) = \lim_{h\to 0} \frac{u(0,h) - u(0,0)}{h} = \lim_{h\to 0} \frac{(-h^3/h^2)}{h} = -1$$

and similarly $v_x(0,0) = 1 = v_y(0,0)$. Thus we see that the Cauchy–Riemann equations are certainly satisfied at the origin. But f is not differentiable at the origin, because for $h = h_1 + ih_1$, $h_1 \in \mathbb{R}$,

$$\frac{f(0+h) - f(0)}{h} = \frac{ih_1}{h_1 + ih_1} = \frac{i}{1+i} = \frac{1+i}{2}$$

and for $h = h_1 + i0$,

$$\frac{f(0+h) - f(0)}{h} = \frac{h_1 + ih_1}{h_1} = 1 + i.$$

Therefore

$$\lim_{h\to 0} \frac{f(0+h) - f(0)}{h}$$

does not exist. This shows that the partial derivatives exist and satisfy the Cauchy–Riemann equations at the origin even though the function is not differentiable there.

A similar conclusion continues to hold for the following function:

$$f(z) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{if } z \neq 0, \\ 0 & \text{if } z = 0. \end{cases}$$

This function vanishes on both coordinate axes. Hence at $z = 0$,

$$u_x = u_y = v_x = v_y = 0,$$

and the Cauchy–Riemann equations are satisfied. But on the line $y = mx$ $(m \neq 0)$, we have

$$\frac{f(h + imh) - f(0)}{h + imh} = \frac{h \cdot mh/(h^2 + m^2h^2)}{h + imh} = \frac{m}{(1+m^2)(1+im)h},$$

which approaches ∞ as h approaches 0. Therefore, $f'(0)$ does not exist. In fact, despite the existence of partial derivatives the function is not even continuous at $z = 0$.

However, even if a function is continuous and has partial derivatives that satisfy the Cauchy–Riemann equations at a point, we still are not guaranteed differentiability. Consider the function

$$f(z) = \begin{cases} \dfrac{xy^2}{x^2+y^2} & \text{if } z \neq 0, \\ 0 & \text{if } z = 0. \end{cases}$$

Since $|f(z)| \leq |x|$, the function is continuous at the origin. Moreover, as $f = 0$ on both the axes, $f_x = f_y = 0$ at the origin, so that the Cauchy–Riemann equations are satisfied at the origin. But on the line $y = mx$ $(m \neq 0)$,

$$\frac{f(h+imh)-f(0)}{h+imh} = \frac{m^2h^3/(1+m^2)h^2}{h+imh} = \frac{m^2}{(1+im)(1+m^2)},$$

which does not tend to a unique limit independent of m. Hence, once again, $f'(0)$ does not exist, meaning that f is not differentiable at the origin. It follows that the converse of Theorem 5.7 is not true.

Thus far, the negative character of the Cauchy–Riemann equations has been emphasized. They have been utilized primarily to prove the *nonexistence* of a derivative. In the next section, we will use these equations in conjunction with an additional criterion to formulate a sufficient condition for differentiability.

As indicated earlier, (5.3) can be expressed more concisely as

$$f'(z) = f_x = -if_y. \tag{5.5}$$

Equation (5.5) provides a method for calculating the derivative if the derivative is known to exist. Note that $z = x + iy$ gives that

$$x = \frac{z+\overline{z}}{2}, \quad y = \frac{z-\overline{z}}{2i}.$$

In view of this note, we may treat $f(x+iy)$ as a function of z and $\overline{z}$, and so

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

Thus

$$\begin{aligned} f_{\overline{z}} = \frac{\partial f}{\partial \overline{z}} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \overline{z}} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \overline{z}} \\ &= \frac{1}{2}\left[\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right] \\ &= \frac{1}{2}\left[(u_x - v_y) + i(v_x + u_y)\right] \end{aligned}$$

and the equation $f_{\overline{z}} = 0$ is equivalent to the system

$$f_x = -if_y, \quad \text{or} \quad u_x = v_y \quad \text{and} \quad u_y = -v_x.$$

Thus, we have

Theorem 5.8. *A necessary condition for a function f to be differentiable at a point a is that it satisfies the equation $f_{\overline{z}} = 0$ at a.*

The differential equation $f_{\overline{z}} = 0$ is known as the homogeneous Cauchy–Riemann equation or simply the complex form of Cauchy–Riemann equations. For instance if $f(z) = \overline{z}$ then $f_{\overline{z}} = 1 \neq 0$, and so it is nowhere differentiable. Similarly, using this we conclude that

$$f_1(z) = \operatorname{Re} z = \frac{z + \overline{z}}{2}, \quad f_2(z) = \operatorname{Im} z = \frac{z - \overline{z}}{2i} \quad \text{and} \quad f_3(z) = e^{\overline{z}}$$

are all nowhere differentiable functions. We now prove an analog of a well-known real-variable theorem.

Theorem 5.9. *If $f'(z) \equiv 0$ in a domain D, then $f(z)$ is constant in D.*

Proof. In view of (5.5), $f_x = f_y = 0$ for all points in D. Now if the derivative of a function of one real variable vanishes in an interval, the function must be constant in that interval. Hence, $u_x = u_y = 0$ in D implies that $u(x, y)$ is constant along every horizontal and vertical line segment in D. Similarly, $v_x = v_y = 0$ in D implies that $v(x, y)$ is constant along every horizontal and vertical line segment in D. Thus, $f(z) = u(z) + iv(z)$ is constant along every polygonal line in D whose sides are parallel to the coordinate axes. According to Remark 2.7, any two points in D can be joined by such a line. Therefore, $f(z_1) = f(z_2)$ for any pair of points $z_1, z_2 \in D$, so that $f(z)$ must be constant in D. ∎

Theorem 5.9 is not true if D in this statement is an open set which is not connected. For example, if $D = \{z : |z| > 2\} \cup \{z : |z| < 1\}$ and

$$f(z) = \begin{cases} 2 & \text{if } |z| < 1 \\ 3 & \text{if } |z| > 3, \end{cases}$$

then $f'(z) = 0$ on D but f is not constant on D.

Theorem 5.10. *If $f(z)$ is real-valued and differentiable on a domain D, then f is constant on D.*

Proof. Suppose that f is differentiable at z_0. Then

$$f'(z_0) = \lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h}.$$

Allowing $h \to 0$ along both the real and imaginary axes, we see that $f'(z_0)$ is both real and pure imaginary. Consequently, $f'(z_0) = 0$. By Theorem 5.9, f is a constant. ∎

An immediate consequence of Theorem 5.9 is the following result.

Corollary 5.11. *If $f = u + iv$ is differentiable in a domain D with either $u(x, y)$, $v(x, y)$, or $\arg f(z)$ constant in D, then $f(z)$ is constant in D.*

Proof. Suppose that $u(x, y) = c$ on D. Then $u_x = u_y = 0$ on D, and by the Cauchy–Riemann equations, $v_y = v_x = 0$. Therefore, v is constant on D. The other cases can be handled by a similar argument. ∎

Corollary 5.12. *If f and g are two differentiable functions in a domain D, and $\operatorname{Re} f(z) = \operatorname{Re} g(z)$ on D, then $f(z) = g(z) + \text{constant}$.*

This corollary says that *"the real part* $\operatorname{Re} f(z)$, *completely determines the differentiable function* $f(z)$ *in a domain except for an additive constant"*. Similarly, every differentiable function f in a domain D is completely determined by its imaginary part except for an additive real constant.

Questions 5.13.

1. What is the geometric interpretation of a continuous function of two variables?
2. For a function of two variables, why is it usually easier to show that a limit does not exist?
3. Why did the behavior of functions along the coordinate axes play a central role?
4. Is there a function f having every directional derivative at a point without f being continuous at that point?
5. Can the Cauchy–Riemann equations be written in polar coordinate form?
6. Can we give a geometric interpretation to the existence of a derivative of a complex function?
7. If a derivative exists, is the derivative a continuous function?
8. When will it be easier to evaluate limits by use of polar coordinates?
9. In this section, we used the word "path" without defining it. How would you define "path"?
10. If a function $f = u + iv$ is differentiable, can we say anything about u_{xx}?
11. If $f(z)$ assumes only real values in a domain, what can we say about the existence of $f'(z)$?
12. Where is $|z|$ differentiable? Where is $|z|^2$ differentiable?
13. Where is $z|z|$ differentiable? Where is $z|z|$ continuous?
14. Where is $\operatorname{Re} z$ differentiable? Where is $(\operatorname{Re} z)^2$ differentiable?
15. Where is $z\operatorname{Re} z$ differentiable? Where is $z\operatorname{Re} z$ continuous?

Exercises 5.14.

1. If $f(0,0) = 0$, which of the following functions are continuous at the origin?

   (a) $f(x,y) = \dfrac{x^2y^2}{x^4+y^4}$ (b) $f(x,y) = \dfrac{x^2y^2}{(x^2+y^2)^2}$

   (c) $f(x,y) = \dfrac{x^3y^2}{(x^2+y^2)^2}$ (d) $f(x,y) = \dfrac{x+ye^{-x^2}}{1+y^2}$

   (e) $f(x,y) = \dfrac{(x+y^2)^2}{x^2+y^2}$ (f) $f(x,y) = \dfrac{x\sqrt{|xy|}}{x^2+y^2}$.
2. Determine where the following functions satisfy the Cauchy–Riemann equations, and where the functions are differentiable.

(a) $f(z) = \overline{z}^2$ (b) $f(z) = x^2 - y^2$
(c) $f(z) = 2xyi$ (d) $f(z) = x^2 - y^2 + 2xyi$
(e) $f(z) = z\mathrm{Re}\, z$ (f) $f(z) = z|z|$
(g) $f(z) = |\mathrm{Re}\, z\, \mathrm{Im}\, z|^{1/2}$ (h) $f(z) = |\mathrm{Re}\, z\, \mathrm{Im}\, z|^{1/3}$
(i) $f(z) = z\mathrm{Im}\, z$ (j) $f(z) = 2z + 4\overline{z} + 5$.

3. Show that at $z = 0$ the function f defined by

$$f(x+iy) = \begin{cases} \dfrac{(1+i)x^3 - (1-i)y^3}{x^2+y^2} & \text{for } x+iy \neq 0 \\ 0 & \text{for } x = y = 0 \end{cases}$$

satisfies the Cauchy–Riemann equations but it is not differentiable.

4. If $f(z)$ is continuous at a point z_0, show that $\overline{f(z)}$ is also continuous at z_0. Is the same true for differentiability at z_0? What does the function $f(z) = |z|^2$ show? How about $f(z) = z$?
5. If $f(z)$ is continuous at a point z_0, then show that $\overline{f(\overline{z})}$ is also continuous at z_0. Is the same is true for the differentiability at z_0?
6. Is $f : \mathbb{C} \to \mathbb{C}$ given by $f(z) = z^2 + z|z|^2$ differentiable at $z = 0$? If so, what is $f'(0)$? Does $f^{(n)}(0)$ exist for $n \geq 2$? Give your reasons.
7. a) Give an integer n and nonzero real number m, construct a function $f(z)$ such that $\lim_{z\to 0} f(z) = 0$ along each curve of the form $y = mx^k$ $(k = 0, 1, 2, \ \ldots \ , n-1)$, but for which $\lim_{z\to 0} f(z) \neq 0$ along the curve $y = mx^n$.
   b) Construct a function $f(z)$ for which $\lim_{z\to 0} f(z) = 0$ along each curve of the form $y = mx^n$ $(n = 1, 2, 3, \ldots)$, but for which $\lim_{z\to 0} f(z)$ does not exist.
8. If $f(z)$ is differentiable at z, show that

$$\begin{aligned} |f'(z)|^2 &= \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 = \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \\ &= \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x} = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} = \frac{\partial(u,v)}{\partial(x,y)}. \end{aligned}$$

The last expression is the Jacobian of u and v with respect to the variables x and y.

9. If f_x and f_y exist at a point (x_0, y_0), show that

$$\lim_{x\to x_0} f(x, y_0) = \lim_{y\to y_0} f(x_0, y) = f(x_0, y_0).$$

Must $f(x, y)$ be continuous at (x_0, y_0)?

5.2 Analyticity

Try as we did, we were unable to extract differentiability from the Cauchy–Riemann equations. This "smoothness" requirement along the coordinate axes could not sufficiently control the behavior of a function along different paths. If we focus on neighborhoods rather than on isolated points, many of our difficulties will be eliminated.

A function is said to be *analytic* at a point if it is differentiable everywhere in some neighborhood of the point. A function is analytic in a domain if it is analytic at every point in the domain. A function analytic at every point in the complex plane is called an *entire* function.

The function
$$f(z) = |z|^2 = x^2 + y^2$$
is differentiable only at the origin, and hence is not analytic anywhere. The function $f(z) = x^2y^2$ is differentiable at all points on each of the coordinate axes, but is still nowhere analytic. On the other hand, all the polynomials are entire functions, and $f(z) = z/(1-z)$ is analytic everywhere except at $z = 1$.

Remark 5.15. If $f(z)$ is analytic at z_0, then there exist an $\epsilon > 0$ such that $f(z)$ is differentiable at each point in $N(z_0;\epsilon)$. But for any point $z_1 \in N(z_0;\epsilon)$, we can find a $\delta > 0$ such that $N(z_1;\delta) \subset N(z_0;\epsilon)$. Hence $f(z)$ is also analytic at z_1. Consequently, $f(z)$ is analytic at a point if and only if $f(z)$ is analytic in some neighborhood of the point. Thus, the set of all values for which a given function is analytic must form an open set. In particular, if a function is analytic in a closed set, then there is always an open set containing the closed set in which the function is analytic. ●

Returning to functions of two variables, we prove the following mean-value theorem:

Theorem 5.16. *Let $f(x, y)$ be defined in a domain D, with f_x and f_y continuous at all points in D. Given a point $(x, y) \in D$, choose δ so that $(x + h, y + k) \in D$ for all points satisfying $\sqrt{h^2 + k^2} < \delta$. Then*
$$f(x + h, y + k) - f(x, y) = f_x(x, y)h + f_y(x, y)k + \epsilon_1 h + \epsilon_2 k,$$
where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as $h \to 0$ and $k \to 0$.

Proof. We write
$$\begin{aligned} &f(x + h, y + k) - f(x, y) \\ &\qquad = \{f(x + h, y + k) - f(x, y + k)\} + \{f(x, y + k) - f(x, y)\}. \end{aligned} \tag{5.6}$$

Now $f(\xi, y + k)$ may be viewed as a differentiable function of the one real variable ξ, where ξ assumes values between x and $(x + h)$. Applying the mean-value theorem for one real variable, we have
$$f(x + h, y + k) - f(x, y + k) = f_x(x + \theta_1 h, y + k)h \quad (|\theta_1| < 1). \tag{5.7}$$

Similarly,

$$f(x, y+k) - f(x, y) = f_y(x, y+\theta_2 k)k \quad (|\theta_2| < 1). \tag{5.8}$$

By the continuity of f_x and f_y,

$$f_x(x+\theta_1 h, y+k) = f_x(x, y) + \epsilon_1 \quad \text{and} \quad f_y(x, y+\theta_2 k) = f_y(x, y) + \epsilon_2,$$

where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as $h \to 0$ and $k \to 0$. In view of the last equation, we may rewrite (5.7) and (5.8) as

$$\begin{aligned} f(x+h, y+k) - f(x, y+k) &= f_x(x, y)h + \epsilon_1 h \\ f(x, y+k) - f(x, y) &= f_y(x, y)k + \epsilon_2 k. \end{aligned}$$

The result now follows from (5.6). ∎

This mean-value theorem enables us to utilize the Cauchy–Riemann equations to establish sufficient conditions for analyticity.

Theorem 5.17. *Let $f(z) = u(x, y) + iv(x, y)$ be defined in a domain D, and let $u(x, y)$ and $v(x, y)$ have continuous partials that satisfy the Cauchy–Riemann equations*

$$u_x = v_y, \quad v_x = -u_y$$

for all points in D. Then $f(z)$ is analytic in D.

Proof. Set $\Delta z = h + ik$, where h and k are arbitrary real numbers. Given a point $z \in D$, we must show that

$$\lim_{\Delta z \to 0} \frac{f(z+\Delta z) - f(z)}{\Delta z}$$

exists. For h and k sufficiently small, $z+\Delta z$ is in D. Since u and v are assumed to have continuous partials, an application of Theorem 5.1 shows that f_x and f_y must also be continuous. Thus the conditions of Theorem 5.16 are satisfied and we have

$$\frac{f(z+(h+ik)) - f(z)}{h+ik} = \frac{f_x(x, y)h + f_y(x, y)k + \epsilon_1 h + \epsilon_2 k}{h+ik}, \tag{5.9}$$

where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as $h \to 0$ and $k \to 0$. As we saw in (5.5), the Cauchy–Riemann equations may be expressed as $f_x(x, y) = -if_y(x, y)$, or

$$if_x(x, y) = f_y(x, y). \tag{5.10}$$

Substituting (5.10) into (5.9), we obtain

$$\begin{aligned} \frac{f(z+(h+ik)) - f(z)}{h+ik} &= \frac{f_x(x, y)h + if_y(x, y)k + \epsilon_1 h + \epsilon_2 k}{h+ik} \qquad (5.11) \\ &= f_x(x, y) + \epsilon_1 \frac{h}{h+ik} + \epsilon_2 \frac{k}{h+ik}. \end{aligned}$$

But

$$\left|\frac{h}{h+ik}\right| \le 1, \quad \left|\frac{k}{h+ik}\right| \le 1,$$

so that the last two expressions in (5.11) approach zero as h and k approach zero. Therefore,

$$\lim_{h\to 0,\, k\to 0} \frac{f(z+(h+ik))-f(z)}{h+ik} = f_x(x,y).$$

Because no assumptions were made about the manner in which h and k approached zero, the derivative $f'(z)$ exists, with $f'(z) = f_x(x,y)$. Since z was arbitrary, the function is differentiable everywhere in D, and hence is analytic in D.

Observe that we could similarly have shown that $f'(z) = -if_y(x,y)$ for all points in D. ■

It pays, at this point, to extract the important steps of Theorem 5.17 in order to understand more fully the relationship between the Cauchy–Riemann equations and differentiability. Requiring continuity of the partials in a neighborhood allowed us to apply Theorem 5.16 to obtain (5.9). A substitution of the Cauchy–Riemann equations into (5.9) led to (5.11). Analyticity then followed from (5.11).

Had differentiability at a *point* been our main objective, we would have proved a theorem analogous to Theorem 5.17 that required continuity of the partials at only one point. For complex functions, however, analyticity (rather than differentiability at a point) is the important concept.

Example 5.18. Consider $f(z) = u+iv$, where $u = x^2$ and $v = y^2$. Each of the partial derivatives is continuous in $\mathbb{C}$, whereas u and v satisfy the Cauchy–Riemann equations only when $y = x$. Thus f is differentiable at $(1+i)x$, $x \in \mathbb{R}$, and nowhere else. ●

Example 5.19. If $u(x,y) = x^2 + y^2$ and $v(x,y) = xy$ then all the partial derivatives exist and are continuous in $\mathbb{C}$. However $f = u + iv$ is differentiable only at $z = 0$ because the Cauchy–Riemann equations are satisfied only at the origin; so f is nowhere analytic since f is differentiable at $z = 0$ and nowhere else. ●

In Chapter 8, we will show that if $f(z)$ is analytic at a point, then $f(z)$ has derivatives of *all* orders at the point. In particular, the existence of $f'(z)$ tells us that

$$f'(z) = \frac{\partial f}{\partial x} = -i\frac{\partial f}{\partial y}$$

is continuous. In view of Theorem 5.1, the partial derivatives of its real and imaginary components are also continuous.

Therefore, the converse of Theorem 5.17 is also true. That is, a function is analytic in a domain D *if and only if* the function has continuous partials that satisfy the Cauchy–Riemann equations for all points in D.

Remark 5.20. In real analysis, if $f(x)$ is differentiable on (a,b), then the derivative $f'(x)$ need not be differentiable on (a,b) (need not even be continuous as shown by an example below). Clearly, the function

$$f(x) = \frac{3}{4}x^{4/3}$$

is differentiable on $(-1,1)$, whereas $f'(x) = x^{1/3}$ is not differentiable at the origin. ●

Note that differentiability in a neighborhood does not assume the same importance for real as for complex functions. The real-valued function

$$f(x) = x|x|$$

is differentiable for all real values, but does not have a second derivative at the origin. The everywhere-differentiable function in $\mathbb{R}$

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

does not even have a continuous derivative at the origin.

It is now quite simple to establish analyticity for the elementary functions.

Examples 5.21. Let $f(z) = e^z = e^x(\cos y + i \sin y)$. We have

$$u_x = v_y = e^x \cos y \quad \text{and} \quad v_x = -u_y = e^x \sin y.$$

Theorem 5.17 may be applied to deduce that $f(z) = e^z$ is an entire function and that

$$f'(z) = \frac{\partial f}{\partial x} = e^x \cos y + ie^x \sin y = e^z.$$

Similarly, if $f(z) = \sin z = (e^{iz} - e^{-iz})/2i$, then by Theorem 5.6 and Example 5.21,

$$f'(z) = \frac{ie^{iz} + ie^{-iz}}{2i} = \frac{e^{iz} + e^{-iz}}{2} = \cos z.$$

Also, if $f(z) = \cos z = (e^{iz} + e^{iz})/2$, then

$$f'(z) = \frac{ie^{iz} - ie^{-iz}}{2} = \frac{-e^{iz} + e^{-iz}}{2i} = -\sin z.$$ ●

The above examples illustrate a recurrent pattern in the theory of complex variables. The real-valued exponential furnishes us with information about the complex exponential, which then enables us to derive results about the complex trigonometric functions. We leave it for the reader to verify that $\tan z$, $\cot z$, $\sec z$ and $\csc z$ have the "expected" derivatives.

For certain functions, it is easier to express real and imaginary parts in terms of r and θ of $z = re^{i\theta}$,

$$f(z) = u(r, \theta) + iv(r, \theta).$$

For example, if $f(z) = z^5$, then expressing this in terms of x and y is uninviting. On the other hand, De Moivre's theorem quickly gives

$$f(z) = (re^{i\theta})^5 = r^5 \cos 5\theta + ir^5 \sin 5\theta.$$

A sometimes useful expression for the Cauchy–Riemann equations in polar coordinates will now be proved.

Theorem 5.22. *Let $f = u + iv$ be differentiable with continuous partials at a point $z = re^{i\theta}$, $r \neq 0$. Then*

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

In most textbooks, the polar form of the two Cauchy–Riemann equations is either left as an exercise or worked out by using the chain rule for functions of two variables to write u_x, u_y, v_x and v_y in terms of r and θ and then solving/comparing the equations which involve these partial derivatives. Let us provide a simpler proof of Theorem 5.22.

Proof. Recall that $f'(z) = f_x = -if_y$. We have for $z = re^{i\theta} \neq 0$,

$$f_r = \frac{\partial f}{\partial z}\frac{\partial z}{\partial r} = f'(re^{i\theta})e^{i\theta} \quad \text{and} \quad f_\theta = f'(re^{i\theta})ire^{i\theta}, \tag{5.12}$$

or equivalently,

$$f_r = e^{i\theta} f'(re^{i\theta}) \quad \text{and} \quad f_\theta = ire^{i\theta} f'(re^{i\theta}) \tag{5.13}$$

which shows that

$$f_r = -\frac{i}{r} f_\theta, \quad \text{i.e.} \quad u_r + iv_r = -\frac{i}{r}(u_\theta + iv_\theta) \quad (r \neq 0).$$

Taking the real and imaginary parts of this equation yields the desired Cauchy–Riemann equations in polar form. ∎

To demonstrate that the polar form is useful, we consider $f : \mathbb{C} \setminus \{0\} \to \mathbb{C}$ given by

$$f(z) = \operatorname{Log} z := \ln|z| + i\operatorname{Arg} z.$$

If $z = re^{i\theta} \neq 0$, then this equation becomes

$$f(z) = \ln r + i\theta := u(r, \theta) + iv(r, \theta) \quad (-\pi < \theta \leq \pi)$$

and it is now easier to check the polar form of the Cauchy–Riemann equations. As (see (5.13))

$$f'(z) = e^{-i\theta} f_r = -\frac{i}{r} e^{-i\theta} f_\theta, \tag{5.14}$$

it follows that

$$\frac{d}{dz}(\operatorname{Log} z) = \frac{e^{-i\theta}}{r} = \frac{1}{z}, \quad z \in \mathbb{C} \setminus \{x + iy : x \leq 0,\ y = 0\}.$$

Note that v is not continuous at points on the negative real axis, and the partial derivatives are continuous at all points except those points on the negative real axis.

For another illustration, consider

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

It is not easy to use the Cartesian form to verify the analyticity of this function. In terms of polar coordinates, we write $x = r\cos\theta$ and $y = r\sin\theta$ so that

$$f(z) = -\frac{\cos 2\theta}{r^2} + i\frac{\sin 2\theta}{r^2} = -\frac{e^{-2i\theta}}{r^2} \quad (r \neq 0).$$

Thus,

$$f_r = \frac{2}{r^3} e^{-2i\theta} \quad \text{and} \quad f_\theta = \frac{2i}{r^2} e^{-2i\theta}$$

which gives, according to the formula (5.14),

$$f'(z) = e^{-i\theta} f_r = \frac{2}{r^3 e^{3i\theta}} = \frac{2}{z^3}, \quad z \neq 0.$$

Example 5.23. We wish to obtain a branch of $(z^2 - 1)^{1/2}$ that is analytic for $|z| > 1$.

To do this we need to identify a solution $w = f(z)$ that is analytic for $|z| > 1$ and satisfies the condition

$$w^2 = z^2 - 1. \tag{5.15}$$

If we use the principal branch of $(z^2 - 1)^{1/2}$, then

$$w = e^{(1/2)\operatorname{Log}(z^2 - 1)}.$$

Does this function do the job? Note that this function is analytic on the open set $\Omega = \{z : z^2 - 1 \in \mathbb{C} \setminus (-\infty, 0]\}$. Note that a point z fails to belong to Ω iff

$$z^2 - 1 \in (-\infty, 0].$$

To exhibit Ω we just need to remove from $\mathbb{C}$ those complex numbers z for which

$$\operatorname{Re}(z^2 - 1) = x^2 - y^2 - 1 \leq 0 \quad \text{and} \quad 2xy = 0.$$

This gives the set

$$\{x+iy : x=0\} \cup \{x+iy : y=0, |x| \le 1\}.$$

Thus, the branch cut for the principal branch of $(z^2-1)^{1/2}$ is the whole y-axis as well as the real interval $[-1,1]$ of $\mathbb{R}$. So we need to look for an alternate solution w satisfying (5.15). In view of this, we rewrite (5.15) as

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

and consider

$$w = g_1(z) = z(1-1/z^2)^{1/2} = ze^{(1/2)\,\mathrm{Log}\,(1-1/z^2)}$$

or

$$w = g_2(z) = -ze^{(1/2)\,\mathrm{Log}\,(1-1/z^2)},$$

each of which is a solution of (5.15). As $\mathrm{Log}\,(1-z)$ is analytic for $z \in \mathbb{C}\setminus[1,\infty)$, $\mathrm{Log}\,(1-1/z)$ is, in particular, analytic for $|z|>1$. Consequently, g_1 and g_2 are analytic for $|z|>1$. What are their derivatives for $|z|>1$?

Finally, to obtain the branch of $(z^2-1)^{-1/2}$ which are analytic for $|z|<1$, we rewrite (5.15) as

$$w^2 = -(1-z^2) = i^2(1-z^2).$$

This allows us to consider two solutions of (5.15) which are analytic for $|z|<1$, namely,

$$w=\phi_1(z) = ie^{(1/2)\,\mathrm{Log}\,(1-z^2)} \quad \text{or} \quad w=\phi_2(z) = -ie^{(1/2)\,\mathrm{Log}\,(1-z^2)}. \quad \bullet$$

Example 5.24. We wish to determine the largest open set Ω in which the function $\mathrm{Log}\,(1-z^n)$ $(n \in \mathbb{N})$ is analytic.

To do this, we first recall that $g(z) = \mathrm{Log}\, z$ is analytic in $\mathbb{C}\setminus(-\infty,0]$ but not in any larger open set. Also, $g'(z)=1/z$. Set $h(z)=1-z^n$. Then g is analytic in $\mathbb{C}$ minus those points z for which

$$1-z^n \in (-\infty,0], \quad \text{i.e., } z^n \in [1,\infty).$$

For instance, if $n=2$, then

$$1-z^2 \in (-\infty,0] \Longleftrightarrow z \in (-\infty,-1] \cup [1,\infty)$$

and

$$1-z^3 \in (-\infty,0] \Longleftrightarrow z \in R_0 \cup R_1 \cup R_2,$$

where $R_k = \{re^{i2k\pi/3} : r \ge 1\}$ $(k=0,1,2)$ is the ray starting from $e^{i2k\pi/3}$ (a cube root of unity) to ∞. Thus, the largest open set in which $\mathrm{Log}\,(1-z^n)$ is analytic is therefore the set

$$\Omega = \mathbb{C} \setminus \cup_{k=0}^{n-1} S_k$$

where $S_k = \{re^{i2k\pi/n} : r \ge 1\}$, $k=0,1,2,\ldots,n-1$. $\bullet$

Example 5.25. Let f be an entire function. Suppose that $f = u + iv$ has the property that

$$u_y - v_x = -2 \quad \text{for all } z \in \mathbb{C}.$$

What can we say about the function f? Can it be a constant function? Clearly not! Can this be a polynomial of degree > 1? The given condition shows that this is not the case (how?). Let us try to find this function. By the Cauchy–Riemann equation $u_y = -v_x$, the given condition is the same as

$$v_x = 1$$

which, by the fact that $f'(z) = u_x + iv_x$, is equivalent to

$$\operatorname{Im} f'(z) = 1.$$

This observation implies that $f'(z)$ is a constant, say a, so that f has the form

$$f(z) = az + b$$

with $\operatorname{Im} a = 1$ (Alternatively, as $\operatorname{Im} f'(z) = 1$, $h(z)$ is defined by

$$h(z) = e^{if'(z)}$$

is entire and $|h(z)| = e^{-1}$ which is a constant. Thus, h and hence, $f'(z)$ is a constant). ●

We close this section with a theorem requiring only differentiability at a point, although a more general theorem for analytic functions will be proved later.

Theorem 5.26. *Let $f(z)$ and $g(z)$ be differentiable at z_0, with $f(z_0) = g(z_0) = 0$. If $g'(z_0) \neq 0$, then $\lim_{z \to z_0}(f(z)/g(z)) = f'(z_0)/g'(z_0)$.*

Proof. The result is a consequence of the definition of a derivative, for

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

Examples 5.27. (i) Let $f(z) = |z|^2$ and $g(z) = z$. Then

$$\lim_{z\to 0} \frac{f(z)}{g(z)} = \frac{f'(0)}{g'(0)} = \frac{0}{1} = 0.$$

Note that for $z \neq 0$, $f(z)/g(z)$ equals $\overline{z}$.

(ii) For $f(z) = \sin z$, we have

$$\lim_{z\to 0} \frac{f(az)}{f(z)} = \lim_{z\to 0} \frac{\sin az}{\sin z} = a\frac{\cos 0}{\cos 0} = a,$$

where a is any complex number.

(iii) Let $f(z) = 1 - \cos z$ and $g(z) = \sin^2 z$. Here $g'(2n\pi) = 0$ for each $n \in \mathbb{Z}$, but Theorem 5.26 may be avoided by solving directly. Because

$$\sin^2 z = 1 - \cos^2 z = (1 - \cos z)(1 + \cos z),$$

we have for each $n \in \mathbb{Z}$

$$\lim_{z\to 2n\pi} \frac{1-\cos z}{\sin^2 z} = \lim_{z\to 2n\pi} \frac{1-\cos z}{1-\cos^2 z} = \lim_{z\to 2n\pi} \frac{1}{1+\cos z} = \frac{1}{2}. \qquad \bullet$$

Questions 5.28.

1. If we do not require continuity of the partials in Theorem 5.16, what kind of mean-value theorem can we obtain?
2. What important differences are there between differentiable functions and analytic functions?
3. Let f be defined on a domain $D \subseteq \mathbb{C}$ and $a, b \in D$ such that $[a, b] \subseteq D$. Does there exists a point c on this line segment $[a, b]$ such that
$$f(b) - f(a) = f'(c)(b - a)?$$
4. What alternate definitions could we have given for an entire function?
5. What relationships are there between continuity of a function and the continuity of its partials?
6. If $f(z)$ and $\overline{f(z)}$ are both analytic in a domain D, what can be said about f throughout D?
7. Let $f(z)$ be analytic in the unit disk $\Delta = \{z \in \mathbb{C} : |z| < 1\}$. Is $g(z) = \overline{f(\overline{z})}$ analytic in Δ?
8. Is $f(z) = (\overline{z})^3 - 3\overline{z}$ differentiable at ± 1? Is $f(z)$ analytic?
9. Does there exist a function f that is analytic only for $\mathrm{Im} z \geq 2004$ and nowhere else?
10. How might we define a real analytic function?
11. If $f(z_0) = g(z_0) = 0$, $f'(z_0)$ and $g'(z_0)$ exist with $g'(z_0) \neq 0$, does
$$\lim_{z\to z_0} \frac{f(z)}{g(z)} = \lim_{z\to z_0} \frac{f'(z)}{g'(z)}?$$
12. If a function is analytic in a bounded region, is the function bounded?
13. If $f(z)$ satisfies the Cauchy–Riemann equations for all points in the plane, is $f(z)$ an entire function?
14. Why isn't the polar form of the Cauchy–Riemann equations valid at the origin?
15. Are the branch cuts of $(z^2 - 4)^{1/2}$ the whole of imaginary axis and the interval $[-2, 2]$ of the real axis? If so, what is the region of analyticity of the chosen branch?
16. Is $\sqrt{z^2 - 4} = \sqrt{z + 2}\sqrt{z - 2}$? If so, when?
17. What is the branch cut for principal branch of $(1 - z^2)^{1/2}$?
18. Does $iz + (1 - z^2)^{1/2}$ take the value 0 at some point $z \in \mathbb{C}$, regardless of the choice of the square root?

19. What is the derivative of $(\sqrt{z})^3$ at $z = 1 - i$?
20. Where is $\operatorname{Arg} z$ continuous? Where is $(\operatorname{Arg} z)^2$ continuous? Where is $(\operatorname{Arg} z)^3$ continuous?
21. Does there exist a function f that is analytic for $\operatorname{Re}(z) \geq 2006$ and is not analytic anywhere else?

Exercises 5.29.

1. Suppose that $f(x) = x^2 - y^3 + i(x^3 + y^2)$. Does it satisfy the Cauchy–Riemann equations for points on the line $y = x$? If so, can we say that f is differentiable at these points? If so, can we say that f is analytic at these points?
2. Place restrictions on the constants a, b, c so that the following functions are entire:
   (a) $f(z) = x + ay - i(bx + cy)$
   (b) $f(z) = ax^2 - by^2 + icxy$
   (c) $f(z) = e^x \cos ay + ie^x \sin(y + b) + c$
   (d) $f(z) = a(x^2 + y^2) + ibxy + c$.
3. Let
$$f(z) = \begin{cases} e^{-1/z^4} & \text{if } z \neq 0 \\ 0 & \text{if } z = 0. \end{cases}$$
   Show that
   (a) $f(z)$ satisfies the Cauchy–Riemann equations everywhere
   (b) $f(z)$ is analytic everywhere except at $z = 0$
   (c) $f(z)$ is not continuous at $z = 0$.
4. Let $f(z)$ and $g(z)$ be entire functions. Show that the following functions are also entire.
$$\text{(a) } f(z) + g(z) \qquad \text{(b) } f(z)g(z) \qquad \text{(c) } f(g(z)).$$
5. Let $f(z)$ and $g(z)$ be analytic at z_0. Show that $f(z)/g(z)$ is analytic at z_0 if and only if $g(z_0) \neq 0$.
6. Let $f(z) = a_0 + a_1 z + \cdots + a_n z^n$. Prove that $a_k = (f^{(k)}(0))/k!$ $(k = 0, 1, \ldots, n)$.
7. If $f(z)$ is analytic at z_0, show that $g(z) = \operatorname{Re} f(z)$ is analytic at z_0 if and only if $g(z)$ is constant in some neighborhood of z_0.
8. Using the Cauchy–Riemann equations, find the most general entire functions such that $\operatorname{Re} f'(z) = 0$.
9. If f is entire such that $f(z_1 + z_2) = f(z_1) + f(z_2)$ for all $z_1, z_2 \in \mathbb{C}$, then show that $f(z) = f(1)z$. Give an example of a continuous nowhere differentiable function f satisfying the given condition for all $z_1, z_2 \in \mathbb{C}$. Find also all continuous functions satisfying this condition.
10. Find the most general entire functions of the form $f(z) = u(x) + iv(y)$, where u and v depend only one real variable.
11. If $f = u + iv$ is differentiable at a point z, show that the first-order partials of u and v exist at z.

12. If f is analytic in a domain D, prove that
   (a) $\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)|f(z)|^n = n^2|f(z)|^{n-2}|f'(z)|^2$
   (b) $\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)|\operatorname{Re} f(z)|^n = n(n-1)|\operatorname{Re} f(z)|^{n-2}|f'(z)|^2.$
13. (a) Let $f(z)$ be analytic with continuous partials in a domain D that excludes the origin. Use Theorem 5.22 to show that

$$f'(z) = e^{-i\theta}\frac{\partial f}{\partial r} = e^{-i\theta}(u_r + iv_r) = \frac{1}{iz}\frac{\partial f}{\partial \theta} = \frac{e^{-i\theta}}{r}(v_\theta - iu_\theta)$$

   at all points in D.
   (b) Conversely, if $f(z)$ has continuous partials in a domain D that excludes the origin, and

$$e^{-i\theta}\frac{\partial f}{\partial r} = \frac{1}{iz}\frac{\partial f}{\partial \theta}$$

   at all points in D, show that $f(z)$ is analytic in D.
14. Use the above exercise to determine whether the following functions are analytic in the domain of definition. What is its derivative in terms of z?
   (a) $r + i\theta$ (b) $\theta + ir$
   (c) $\ln r + i(\theta + 2\pi)$ (d) $r^{10}\cos(10\theta)$
   (e) $r^{10}\sin(10\theta)$ (f) $\dfrac{10r^2 - \sin(2\theta)}{r^2}$.
   where $0 < r$ and $-\pi < \theta < \pi$.
15. Use the result of the previous exercise to show that the derivative of z^n, n an integer, is nz^{n-1}.
16. Evaluate the following limits, if they exist.
   (a) $\displaystyle\lim_{z\to 0}\frac{e^z - 1}{3z}$ (b) $\displaystyle\lim_{z\to 0}\frac{z^2}{|z|}$
   (c) $\displaystyle\lim_{z\to 0}\frac{2\sin z}{e^z - 1}$ (d) $\displaystyle\lim_{z\to 0} z\sin\frac{1}{z}$.
17. Construct a branch $f(z)$ of $\log z$ such that $f(z)$ is analytic at $z = -1$ and takes on the value $5\pi i$ there.

5.3 Harmonic Functions

If $f(z)$ is analytic in a domain, then its derivative can be expressed in several forms. For instance, according to (5.5) we may write

$$f'(z) = \frac{\partial f}{\partial x}, \quad \text{or} \quad f'(z) = -i\frac{\partial f}{\partial y}. \tag{5.16}$$

In Chapter 8 it will be shown that an analytic function has derivatives of all orders. From the first equation in (5.16), we see that the second derivative can be expressed as

$$f''(z) = \frac{\partial}{\partial x} f'(z) = \frac{\partial^2 f}{\partial x^2}. \tag{5.17}$$

In view of the second equation in (5.16), we also have

$$f''(z) = -i\frac{\partial}{\partial y} f'(z) = -i\frac{\partial}{\partial y}\left(-i\frac{\partial f}{\partial y}\right) = -\frac{\partial^2 f}{\partial y^2}. \tag{5.18}$$

Equating (5.17) and (5.18), we get the identity

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0, \tag{5.19}$$

which is valid for any analytic function $f(z)$. Thus if $f(z) = u(x,y) + iv(x,y)$ is analytic in a domain, equation (5.19) shows that its real and imaginary components must satisfy the partial differential equations

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad \text{and} \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0. \tag{5.20}$$

A continuous real-valued function $U(x,y)$, defined in a domain D, is said to be *harmonic* in D if it has continuous first- and second-order partials that satisfy the equation

$$U_{xx} + U_{yy} = 0, \tag{5.21}$$

known as *Laplace's equation*, throughout D. Thus, in the case of the functions of two variables, the above discussion provides the intimate connection between analytic functions and harmonic functions in the following form.

Theorem 5.30. *If $f = u + iv$ is analytic on a domain D, and the functions u and v have continuous second order partial derivatives on D, then u and v are harmonic on D.*

Using the (as yet unproved) result that a function analytic in a domain has derivatives of all orders at each point in the domain (hence a continuous second derivative), we see from (5.20) the following result which is a reformulation of Theorem 5.30.

Theorem 5.31. *Both the real and imaginary parts of an analytic function are harmonic.*

Let us now obtain the polar form of Laplace's equation. From (5.13), we have

$$
\begin{aligned}
f_{rr} &= \frac{\partial}{\partial r}(e^{i\theta} f'(re^{i\theta})) = e^{i\theta}[e^{i\theta} f''(re^{i\theta})] = e^{2i\theta} f''(re^{i\theta}),\\
f_{\theta\theta} &= \frac{\partial}{\partial \theta}(ire^{i\theta} f'(re^{i\theta}))\\
&= ire^{i\theta}\frac{\partial}{\partial\theta}(f'(re^{i\theta})) + i^2 re^{i\theta} f'(re^{i\theta})\\
&= ire^{i\theta}[ire^{i\theta} f''(re^{i\theta})] - re^{i\theta} f'(re^{i\theta})\\
&= -r^2 e^{2i\theta} f''(re^{i\theta}) - rf_r\\
&= -r^2 f_{rr} - rf_r.
\end{aligned}
$$

We thus obtain

$$
\frac{\partial^2 f}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{1}{r}\frac{\partial f}{\partial r} = 0, \quad \text{i.e.} \quad r\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{\partial^2 f}{\partial \theta^2} = 0
$$

which is the polar form for Laplace's equation.

If u is harmonic on D such that $f(z) = u(x,y) + iv(x,y)$ is analytic, then v is called a *harmonic conjugate* of u.

Remark 5.32. We have the antisymmetric property that v is a harmonic conjugate of u if and only if u is harmonic conjugate of $-v$. This follows upon observing that the function $if = i(u+iv) = -v+iu$ is analytic whenever f is analytic. ●

Although $f(z) = x+iy$ is analytic so that $v(x,y) = y$ is harmonic conjugate of $u(x,y) = x$, $g(z) = v+iu = i(x-iy) = i\overline{z}$ is nowhere analytic. This example is to illustrate the following: *"If v is a harmonic conjugate of u in some domain D, then u is not a harmonic conjugate of v unless $u+iv$ is a constant"*.

Laplace's equation furnishes us with a necessary condition for a function to be the real (or imaginary) part of an analytic function.

Example 5.33. For the function $u(x,y) = x^2 + y$, we have

$$
u_{xx} + u_{yy} \equiv 2,
$$

so that u satisfies Laplace's equation nowhere. Hence, $u(x,y) = x^2 + y$ cannot be the real part of *any* analytic function. ●

We will now show how the Cauchy–Riemann equations may be applied to find a harmonic conjugate. For instance, the function $u(x,y) = x + e^{-x}\cos y$ can easily be shown to be harmonic everywhere. If there exists a function $v(x,y)$ for which $f(z) = u(x,y) + iv(x,y)$ is analytic in $\mathbb{C}$, then

$$
u_x = 1 - e^{-x}\cos y = v_y. \tag{5.22}
$$

Antidifferentiation of (5.22) with respect to y yields

$$
v = y - e^{-x}\sin y + \phi(x), \tag{5.23}
$$

where $\phi(x)$ is a differentiable function of x. But in view of (5.23), an application of the other Cauchy–Riemann equation leads to

$$u_y = -e^{-x}\sin y = -v_x = -e^{-x}\sin y - \phi'(x),$$

which can be valid only if $\phi'(x) \equiv 0$. Hence,

$$v(x,y) = y - e^{-x}\sin y + c,$$

where c is a real constant. Therefore, $v(x,y)$ is a harmonic conjugate of $u(x,y)$ and, according to Theorem 5.17, $f(z)$ is an entire function. In fact,

$$\begin{aligned} f(z) &= x + e^{-x}\cos y + i(y - e^{-x}\sin y + c) \\ &= x + iy + e^{-x}(\cos y - i\sin y) + ic \\ &= z + e^{-z} + ic. \end{aligned}$$

Example 5.34. Suppose we wish to find all analytic functions $f(z)$ whose real part is $u(x,y) = e^x(x\cos y - y\sin y)$. We may simply rewrite u as

$$u(x,y) = e^x \mathrm{Re}\,((x+iy)(\cos y + i\sin y)) = \mathrm{Re}\,[(x+iy)e^x e^{iy}] = \mathrm{Re}\,[ze^z],$$

and so the desired analytic functions are of the form $f(z) = ze^z + ic$ for some $c \in \mathbb{R}$.

Similarly, we may rewrite

$$\begin{aligned} e^{-x}(x\sin y - y\cos y) &= e^{-x}\mathrm{Re}\,[(x+iy)(\sin y + i\cos y)] \\ &= \mathrm{Re}\,[e^{-x}(x+iy)i(\cos y - i\sin y)] \\ &= \mathrm{Re}\,[i(x+iy)e^{-(x+iy)}] \\ &= \mathrm{Re}\,[ize^{-z}]. \end{aligned}$$

Thus, every analytic function, whose real part is

$$u(x,y) = e^{-x}(x\sin y - y\cos y),$$

must be of the form $f(z) = ize^{-z} + ic$ for some real constant c.

In this way one can write

$$\begin{aligned} e^x[(x^2-y^2)\cos y - 2xy\sin y] &= e^x\mathrm{Re}\,[(x^2 - y^2 + 2ixy)(\cos y + i\sin y)] \\ &= \mathrm{Re}\,[z^2 e^x(\cos y + i\sin y)] \\ &= \mathrm{Re}\,[z^2 e^z], \end{aligned}$$

and conclude that every analytic function, whose real part is given by

$$u(x,y) = e^x[(x^2-y^2)\cos y - 2xy\sin y],$$

must be of the form $z^2 e^z + ic$ for some real c. •

Example 5.35. Suppose that $f(z) = u + iv$ is an analytic function in a domain D such that $u - v$ is given as

$$u - v = e^x(\cos y - \sin y).$$

We wish to find $f(z)$ in terms of z. The easy procedure is as follows. As $f = u + iv$, we have $if = -v + iu$ so that

$$\mathrm{Re}\,[(1+i)f] = u - v$$

and therefore, we may write the given expression for $u - v$ as

$$\begin{aligned}\mathrm{Re}\,((1+i)f(z)) &= e^x(\cos y - \sin y)\\ &= e^x \mathrm{Re}\,[(1+i)(\cos y + i\sin y)]\\ &= \mathrm{Re}\,[(1+i)e^x e^{iy}]\\ &= \mathrm{Re}\,[(1+i)e^z + ic] \quad (c \in \mathbb{R})\end{aligned}$$

which shows that

$$f(z) = e^z + i(c/(1+i))$$

for some real constant c. The same procedure may be adopted for problems similar to this. ●

Example 5.36. Let us determine the entire function $f = u + iv$ for which $f(0) = i$ and $u(x,y) = x^4 + y^4 - 6x^2y^2 - 4xy$. To do this, it suffices to compute

$$u_x = 4x^3 - 12xy^2 - 4y \quad\text{and}\quad u_y = 4y^3 - 12x^2y - 4x.$$

Clearly, u is harmonic in $\mathbb{C}$. Now, we have

$$f'(z) = u_x - iu_y = 4x^3 - 12xy^2 - 4y + i(-4y^3 + 12x^2y + 4x) = 4z^3 + 4iz$$

which gives

$$f(z) = z^4 + 2iz^2 + c.$$

Setting $c = i$, we get the desired function. ●

Two questions now arise: To what extent is Laplace's equation *sufficient* to guarantee the existence of a harmonic conjugate, and how do we, in general, determine all such conjugate functions? Both of these questions will be answered in Chapter 10 when we construct a harmonic conjugate for any function harmonic in some neighborhood of a point. Thus, a function will be shown to be harmonic in a neighborhood of a point if and only if it is the real part of some analytic function.

As might be expected, any two harmonic conjugates of a given harmonic function differ by a real constant. For, if $v(x,y)$ and $v^*(x,y)$ are harmonic conjugates of $u(x,y)$, then both $u + iv$ and $u + iv^*$ are analytic so that the difference $i(v - v^*)$ is also analytic. Consequently, $v - v^*$ is a real-valued analytic function. Thus,

$$v(x, y) = v^*(x, y) + C,$$

where C is a constant.

Many properties of analytic functions are inherited from their real or imaginary parts. For example, if $u = \operatorname{Re} f(z)$ is constant in a region where $f(z)$ is analytic, then $f(z)$ must be constant. This follows on applying the Cauchy–Riemann equations to obtain $u_x = u_y = v_x = v_y = 0$.

While the real and imaginary parts of an analytic function are harmonic, its modulus need not be harmonic. However, properties of the analytic function may still be deduced by studying the behavior of its modulus.

Theorem 5.37. *Let $|f(z)|$ be constant in a domain D where $f(z)$ is analytic. Then $f(z)$ is constant in D.*

Proof. If $|f(z)| = |u + iv| = C$, then $u^2 + v^2 = C^2$. Differentiating, we have

$$uu_x + vv_x = 0, \quad uu_y + vv_y = 0. \tag{5.24}$$

An application of Cauchy–Riemann equations to (5.24) yields

$$uu_x - vu_y = 0, \quad uu_y + vu_x = 0. \tag{5.25}$$

Eliminating u_y from (5.25), we get

$$(u^2 + v^2)u_x = 0,$$

so that $u_x = 0$. In a similar manner, we can show that $u_y = v_x = v_y = 0$. Thus, we observe that

$$0 = f'(z) = u_x + iu_y$$

which gives that u and v are constants. ∎

Theorem 5.37 is actually guided through the start of the proof of Theorem 5.9. Here is an alternate proof of Theorem 5.37: Suppose that $|f(z)|^2 = c$ for all $z \in D$ and for some constant $c \in \mathbb{R}$. If $c = 0$, then $|f(z)|^2 = 0$ implies that f vanishes on D. If $c \neq 0$, then $f(z) \neq 0$ on D and $1/f(z)$ is analytic on D. But then

$$c = |f(z)|^2 = f(z)\overline{f(z)}$$

shows that $\overline{f(z)} = c/f(z)$ is analytic on D. Consequently, the real-valued function (see Theorem 5.10)

$$\operatorname{Re} f(z) = \frac{f(z) + \overline{f(z)}}{2}$$

is analytic on D and so it is constant. Similarly,

$$\operatorname{Im} f(z) = \frac{f(z) - \overline{f(z)}}{2i}$$

is a real-valued analytic function on D and so it is a constant. Therefore

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

is constant. Thus, we complete the proof of Theorem 5.37.

The word "domain" in Theorem 5.37 cannot be replaced by "circle". To see this, observe that the nonconstant entire function $f(z) = z^n$, n a positive integer, has constant modulus on any circle centered at the origin.

Thus far, our discussion of analyticity has been confined to single-valued functions. It makes no sense for us to talk about derivatives of multiple-valued functions; because in considering the expression

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0},$$

we have no consistent rule to determine which value for $f(z)$ to take as z varies towards z_0. It does, however, make sense to discuss the analyticity for a fixed branch of a multiple-valued function. Recall that

$$f(z) = \operatorname{Log} z = \ln |z| + i\operatorname{Arg} z \quad (z \neq 0, -\pi < \operatorname{Arg} z < \pi)$$

is a single-valued function, continuous when $-\pi < \operatorname{Arg} z < \pi$. Since $f(z)$ is not continuous on the negative real axis, it certainly is not differentiable there.

We will now show that $f(z)$ is analytic at all points of continuity. Switching to the polar representation, we get

$$f(z) = \operatorname{Log} z = \ln r + i\theta \quad (z = re^{i\theta}, -\pi < \theta < \pi). \tag{5.26}$$

Note that

$$\frac{\partial f}{\partial r} = \frac{1}{r} \text{ and } \frac{\partial f}{\partial \theta} = i.$$

In view of Exercise 5.29(13), it follows that $f'(z)$ exists, with

$$f'(z) = e^{-i\theta}\frac{\partial f}{\partial r} = \frac{1}{iz}\frac{\partial f}{\partial \theta} = \frac{1}{z} \quad (-\pi < \operatorname{Arg} z < \pi).$$

Alternatively (since $w = \log z \iff z = e^w = e^{\operatorname{Log} z + 2k\pi i}$), we simply use the chain rule to differentiate

$$z = e^{\operatorname{Log} z}, \quad z \in D = \mathbb{C} \backslash \{x + i0 : x \leq 0\}$$

and obtain

$$1 = e^{\operatorname{Log} z}\frac{d}{dz}(\operatorname{Log} z) = z\frac{d}{dz}(\operatorname{Log} z).$$

Thus,

$$\frac{d}{dz}(\operatorname{Log} z) = \frac{1}{z}, \quad z \in D.$$

Any other branch of $\log z$, with branch cut along the negative real axis, differs from $\operatorname{Log} z$ in the cut plane by a multiple of $2\pi i$, and hence also has the same derivative $1/z$ in the cut plane.

While $\log z$ is discontinuous for negative real values when the branch cut is along the negative real axis, we can find a branch of $\log z$ for which $\log z$ is analytic on the negative real axis. To illustrate, we can easily show that the function

$$f(z) = \log z \quad (0 < \operatorname{Arg} z < 2\pi)$$

is analytic at all points in the plane cut along the *positive* real axis, with $f'(z) = 1/z$ at all such points. Note that this function cannot be defined to be continuous for positive real values. Thus, given any nonzero complex number z_0, there exists a branch for which $\log z$ is analytic at z_0 with derivative $1/z_0$.

As we have seen in the previous chapter, determining properties of the logarithm enables us to determine properties of several related classes of functions. To illustrate, we can choose a definite branch of the logarithm, and write

$$f(z) = z^{1/2} = e^{(1/2)\log z} \quad (z \neq 0).$$

Then by the chain rule,

$$f'(z) = \frac{1}{2z} e^{\frac{1}{2}\log z} = \frac{1}{2e^{\log z}} e^{\frac{1}{2}\log z} = \frac{1}{2e^{\frac{1}{2}\log z}} = \frac{1}{2z^{1/2}}$$

where we have used the same branch of $z^{1/2}$ on both sides of the identities. More generally, if $f(z) = z^{\alpha} = e^{\alpha \log z}$ for some complex number α and some determination of $\log z$ for z on the cut plane (which depends on our choice), then

$$f'(z) = \frac{\alpha}{z} e^{\alpha \log z} = \alpha z^{\alpha - 1}.$$

Example 5.38. We wish to determine the largest domain D in which the principal branch of $\sqrt{e^z + 1}$ is analytic, and compute its derivative for z in that domain. To do this, we recall that the principal branch of $\sqrt{e^z + 1}$ is

$$f(z) = e^{(1/2)\operatorname{Log}(e^z + 1)}.$$

Note that $\operatorname{Log} z$ is analytic in $\mathbb{C} \setminus (-\infty, 0]$, but not in any larger domain. Consequently, the largest domain of analyticity is $\mathbb{C}$ minus those points in the complex plane for which $e^z + 1$ is real and ≤ 0. To find these points, we note that

- $e^z = e^x e^{iy}$ is real iff $y = n\pi$ for some $n \in \mathbb{Z}$
- $e^z > 0$ if n is even
- $e^z < 0$ if n is odd.

For even n, and $z = n\pi$, we have $e^z = e^x > 0$ for all x and so

$$e^z + 1 = e^x + 1 > 0 \quad \text{for all} \quad x \in \mathbb{R}.$$

On the other hand, if n is odd and $z = n\pi$, then $e^z = -e^x$ so that

$$e^z + 1 = -e^x + 1 \le 0 \ \text{ iff } \ x \ge 0.$$

Thus, the domain of analyticity is $\mathbb{C} \setminus \{x + iy : \, x \ge 0, \, y = (2k+1)\pi, \, k \in \mathbb{Z}\}$ and

$$f'(z) = \frac{e^z}{2\sqrt{e^z+1}} = \frac{e^{z-(1/2)\,\mathrm{Log}\,(e^z+1)}}{2}. \qquad \bullet$$

Remark 5.39. The functions z^α and α^z should not be confused. The former is a multiple-valued function of z when α is not an integer; each branch is analytic in the cut plane and has derivative equal to $\alpha z^{\alpha-1}$. The latter, α^z, which can be expressed as $e^{z \log \alpha}$, is a single-valued entire function once a branch is chosen for $\log \alpha$ $(\alpha \neq 0)$; the derivative of α^z is then given by $\alpha^z \log \alpha$. $\bullet$

We can also use the chain rule and the logarithm to find the derivatives of the inverse trigonometric functions. Recall that

$$\sin^{-1} z = -i \log[iz + (1-z^2)^{1/2}].$$

Since

$$\frac{d}{dz}(1-z^2)^{1/2} = -z(1-z^2)^{-1/2},$$

where the same branch is used on both sides of the equation, it follows that

$$\frac{d}{dz}(\sin^{-1} z) = \frac{-i[i - z(1-z^2)^{-1/2}]}{iz + (1-z^2)^{1/2}} = \frac{1}{(1-z^2)^{1/2}}.$$

More generally, if f is analytic and nonzero at a point z, then a branch may be chosen for which $\log f$ is also analytic in a neighborhood of z, with

$$\frac{d}{dz}[\log f(z)] = \frac{f'(z)}{f(z)}.$$

Furthermore, $f(z) = |f(z)|e^{i \arg f(z)}$ so that

$$\log f(z) = \log |f(z)| + i \arg f(z).$$

Hence $\log |f(z)|$, since it is the real part of an analytic function, is harmonic at all points where $f(z)$ is analytic and nonzero. Setting $f = u + iv$, the harmonicity of $\log |f(z)| = \log \sqrt{u^2 + v^2}$ may also be proved directly by laborious computation.

Questions 5.40.

1. Under what conditions may we deduce properties of analytic functions from those of harmonic functions?
2. What properties do analytic and harmonic functions not have in common?
3. When will $\arg f(z)$ be a harmonic function?
4. How do the properties of $|f(z)|$ and $\log|f(z)|$ compare?
5. If a function $u(x, y)$ is harmonic everywhere in a domain D, does there exist a function $v(x, y)$ for which $u(x, y)+iv(x, y)$ is analytic everywhere in D?
6. Is $f = u + iv$ analytic on a domain whenever u and v are harmonic on D?
7. Does $f(z) = u_x - iu_y$ represent an analytic function on a domain D whenever $u = u(x, y)$ is harmonic on D?
8. Let u and v be harmonic on a domain D, and $U = u_y - v_x$ and $V = u_x + v_y$ on D. Is $F(z) = U + iV$ analytic on D?
9. What is the largest domain on which $f(z) = z^z$ is analytic?
10. What is the largest domain on which $f(z) = 3^{z^3}$ is analytic?
11. For what values of z does $u = x^3 - y^3$ satisfy the Laplace equation? Why is this function not harmonic?
12. Does there exist an analytic function with real or imaginary part as $y^2 - 2xy$?
13. Suppose that $f(z)$ is analytic for $|z| < 1$ such that $f(1/2) = 3$ and $|f(z)|$ is constant for $|z| < 1$. Does $f(z) = 3$ for $|z| < 1$?

Exercises 5.41.

1. Let

$$f(z) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & \text{if } x^2 + y^2 \neq 0, \\ 0 & \text{if } x^2 + y^2 = 0. \end{cases}$$

   Show that $f_{xy}(0, 0) = (f_x)_y(0, 0) \neq f_{yx}(0, 0) = (f_y)_x(0, 0)$.
2. Show that the following functions are harmonic, and then determine their harmonic conjugates.
   (a) $u = ax + by$, a and b real constants
   (b) $u = \dfrac{y}{x^2 + y^2}$, $x^2 + y^2 \neq 0$
   (c) $u = x^3 - 3xy^2$
   (d) $u = \operatorname{Arg} z$, $-\pi < \operatorname{Arg} z < \pi$
   (e) $u = e^{x^2-y^2} \cos 2xy$
   (f) $u = 2xy + 3x^2y - y^3$.
3. Using the Cauchy–Riemann equation or the idea of Example 5.34 find all the analytic functions $f = u + iv$, where u is given as below:
   (a) $u(x, y) = e^x(x \sin y + y \cos y)$, $\quad (x, y) \in \mathbb{R}^2$

(b) $u(x,y) = \dfrac{x}{x^2+y^2}$, $\quad (x,y) \in \mathbb{R}^2 \backslash (0,0)$
(c) $u(x,y) = \sin x \cosh y$, $\quad (x,y) \in \mathbb{R}^2$
(d) $u(x,y) = x^3 - 3xy^2 - 2x$, $\quad (x,y) \in \mathbb{R}^2$
(e) $u(x,y) = x^2 - y^2 - 2xy + 2x - 3y$, $\quad (x,y) \in \mathbb{R}^2$.

4. Let $f = u + iv$ be an entire function such that
$$u + v = 2e^{-x}(\cos y - \sin y).$$
Construct f in terms of z.
5. For what nonnegative integer values of n is the real-valued function $u(x,y) = x^n - y^n$ harmonic? Find the corresponding harmonic conjugate function in each possible values of n.
6. Show that $u(x,y) = xy$ is harmonic in $\mathbb{R}^2$. Find the conjugate harmonic function $v(x,y)$ in $\mathbb{R}^2$. Write $u + iv$ in terms of z.
7. Choose the constant a so that the function $u = ax^2y - y^3 + xy$ is harmonic, and find all harmonic conjugates.
8. Choose the constants a, b, c so that $u = ax^3 + bx^2y + cxy^2 + dy^3$ gives the more general harmonic polynomial, and find all the harmonic conjugates.
9. Choose the constant a so that $u = ax^3 + xy^2 + x$ is harmonic in $\mathbb{C}$ and find all analytic functions f whose real part is the given u.
10. Show that neither $xy(x-y)$ nor $xy(x-2y)$ can be a real part of an analytic function.
11. Use the method of this section to attempt to find a function $v(x,y)$ for which $x^2 + iv(x,y)$ is analytic, and explain where the method breaks down.
12. If u_1 and u_2 are harmonic on D, show that $au_1 + bu_2$, a and b real constants, is also harmonic at D.
13. Suppose u and v are conjugate harmonic functions. Show that uv is a harmonic function. What are the most general conditions for which the product of two harmonic functions is harmonic?
14. If $f = u + iv$ is analytic on a domain D, then show that uv is harmonic on D.
15. Find all harmonic functions $u(x,y)$ in the unit disk $x^2 + y^2 < 1$ such that $u_y = 0$ for $x^2 + y^2 < 1$. What can be said about $u(x,y)$?
16. Suppose $f(z)$ has a derivative of order n. Show that
$$f^{(n)}(z) = \frac{\partial^{(n)} f}{\partial x^{(n)}} = (-i)^n \frac{\partial^{(n)} f}{\partial y^{(n)}}.$$
17. Show that if the real and imaginary parts of both $f(z)$ and $zf(z)$ are harmonic in a domain D, then $f(z)$ is analytic in D.
18. Define a branch of $(1-z^2)^{1/2}$ so that $f(z) = iz + (1-z^2)^{1/2}$ is analytic in the domain $\Omega = \mathbb{C} \backslash \{z : \operatorname{Im} z = 0 \text{ and } |\operatorname{Re} z| \geq 1\}$.
19. Find the derivative of the following functions:
$$\text{(a) } \cos^{-1} z \qquad \text{(b) } \tan^{-1} z \qquad \text{(c) } \sec^{-1} z.$$

6

Power Series

From sequences of numbers, we turn to sequences of functions. Then our concern is with both the *form* of convergence and the *behavior* of the limit function. Convergence, determined at each point in a set, need not require the limit function to retain any of the properties common to each function in the sequence. But if a certain “rapport” exists between the sequence of functions and the set, then the limit function will be forced to confirm to definite standards established by the sequence. This stronger type of convergence, in which the set takes precedence over its points, is called *uniform convergence.*

The most important sequences of functions are those expressible as power series. The limit functions for this class are always analytic inside their regions of convergence. In many instances, a power series behaves like a “big” polynomial.

6.1 Sequences Revisited

Given a (real or complex) sequence $\{a_n\}_{n\geq 1}$, we associate a new sequence $\{s_n\}_{n\geq 1}$ of partial sums s_n related by

$$s_n = \sum_{k=1}^{n} a_k.$$

The symbol $\sum_{k=1}^{\infty} a_k$ is called a *series.* The series is said to converge or to diverge according as the sequence $\{s_n\}_{n\geq 1}$ is convergent or divergent. If $\{s_n\}_{n\geq 1}$ converges to s, the sum (or value) of the series is said to be s, and we write

$$s = \lim_{n\to\infty} s_n = \sum_{k=1}^{\infty} a_k.$$

Hence for convergent series the same symbol is used to denote both the series and its sum. We call s_n the nth *partial sum* of the series and a_n the nth *term* in the series.

By definition, every theorem about series may be formulated as a theorem about sequences (of partial sums). The converse is also true.

Theorem 6.1. *Given any sequence $\{s_n\}_{n\geq 1}$, there exists a sequence $\{a_n\}_{n\geq 1}$ such that $s_n = \sum_{k=1}^{n} a_k$ for every n.*

Proof. We choose $a_1 = s_1$ and $a_n = s_n - s_{n-1}$ for $n > 1$. Then

$$\sum_{k=1}^{n} a_k = a_1 + \sum_{k=2}^{n}(s_k - s_{k-1}) = s_1 + (s_n - s_1) = s_n. \qquad \blacksquare$$

Thus, the definition of series does not furnish us with a "new" concept. It merely provides an additional way of stating new theorems and restating old ones. Because taking limits of sequences is a linear operation, it follows that summing series is also a linear operation. For instance, if $s_n \to s$ and $t_n \to t$, then

$$s_n \pm t_n \to s \pm t \quad \text{and} \quad cs_n \to cs,$$

where c is a complex constant. If we apply these properties to partial sums of series we conclude the following:

Proposition 6.2. *If $\sum_{k=1}^{\infty} a_k = \alpha$ and $\sum_{k=1}^{\infty} b_k = \beta$, then*

$$\sum_{k=1}^{\infty}(a_k \pm b_k) = \alpha \pm \beta \quad \textit{and} \quad \sum_{k=1}^{\infty} ca_k = c\alpha.$$

The *Cauchy criterion* for sequences (Theorem 2.20) may be reworded as follows: Let $s_n = \sum_{k=1}^{n} a_k$. The series $\sum_{k=1}^{\infty} a_k$ converges if and only if for every $\epsilon > 0$, there exists an integer N such that $m, n > N$ implies

$$|s_m - s_n| = \left|\sum_{k=n+1}^{m} a_k\right| < \epsilon. \tag{6.1}$$

By letting $m = n + p$, (6.1) may be written as

$$|s_{n+p} - s_n| = \left|\sum_{k=n+1}^{n+p} a_k\right| < \epsilon \quad (p = 1, 2, \dots), \quad \text{for} \quad n > N. \tag{6.2}$$

The Cauchy criterion is the most general test for convergence of a series. Some of the methods frequently used in elementary calculus, like the ratio and integral tests, require very restrictive hypotheses, and even then do not supply necessary as well as sufficient conditions for convergence.

Many of the familiar properties of series are immediate consequences of (6.2). For example,

Proposition 6.3. *A necessary condition for the series $\sum_{k=1}^{\infty} a_k$ to be convergent is that $a_k \to 0$ as $k \to \infty$.*

Proof. This result follows on setting $p = 1$ in (6.2). Indeed, if $\sum_{k=1}^{\infty} a_k$ converges, $\{s_n\}$ converges. Therefore, $s_n - s_{n-1} = a_n \to 0$ as $n \to \infty$. ∎

It is often useful to consider along with the given series $\sum_{k=1}^{\infty} a_k$ the corresponding series of moduli $\sum_{k=1}^{\infty} |a_k|$. A series $\sum_{k=1}^{\infty} a_k$ is said to be *absolutely convergent* if $\sum_{k=1}^{\infty} |a_k|$ converges. Applying

$$\left| \sum_{k=n+1}^{n+p} a_k \right| \leq \sum_{k=n+1}^{n+p} |a_k|$$

to (6.2), we see that the absolute convergence of a series guarantees its convergence. More precisely, we have

Proposition 6.4. *If a series $\sum a_k$ converges absolutely, then $\sum a_k$ converges.*

Similarly, if $|a_n| \leq K|b_n|$ for every n and for some $K > 0$, the convergence of $\sum_{n=1}^{\infty} |a_n|$ may be deduced from the convergence of $\sum_{n=1}^{\infty} |b_n|$ by applying (6.2). Proposition 6.4 can be proved directly with the help of a little trick. Express

$$\operatorname{Re} a_k = (\operatorname{Re} a_k + |a_k|) - |a_k|.$$

Since $|\operatorname{Re} a_k| \leq |a_k|$, we have

$$0 \leq \operatorname{Re} a_k + |a_k| \leq 2|a_k|.$$

Hence, $\sum(\operatorname{Re} a_k + |a_k|)$ converges and $\sum \operatorname{Re} a_k$, being a difference between two convergent series, converges. Similarly, $\sum \operatorname{Im} a_k$ converges. Consequently, $\sum a_k$ converges.

Suppose all the terms of the sequence $\{a_n\}$ are real and positive. Then

$$s_n - s_{n-1} = a_n > 0,$$

and the sequence of partial sums $\{s_n\}$ is a monotonically increasing sequence. Since a monotone sequence of real numbers converges if and only if it is bounded (Theorem 2.15), the series $\sum_{k=1}^{\infty} a_k$ of positive real numbers converges if and only if the sequence $\{s_n\}$ is bounded. That the positivity of $\{a_n\}$ cannot be dropped from the hypotheses is seen by letting $a_n = (-1)^n$. Here, the series $\sum_{k=1}^{\infty} a_k$ does not converge despite the fact that

$$s_n = \sum_{k=1}^{n} (-1)^k$$

is bounded in absolute value by 1.

Our next theorem shows an interesting relationship between a sequence and its sequence of partial sums.

Theorem 6.5. *Suppose that $a_n > 0$ for every $n \in \mathbb{N}$ and that $\sum_{n=1}^{\infty} a_n$ diverges. If $s_n = \sum_{k=1}^{n} a_k$, then*

(i) $\sum_{n=1}^{\infty} \dfrac{a_n}{s_n}$ *diverges;*

(ii) $\sum_{n=1}^{\infty} \dfrac{a_n}{s_n^2}$ *converges.*

Proof. According to the Cauchy criterion, the series in (i) will diverge if, for any integer n, an integer p can be found such that

$$\frac{a_{n+1}}{s_{n+1}} + \frac{a_{n+2}}{s_{n+2}} + \cdots + \frac{a_{n+p}}{s_{n+p}} > \frac{1}{2}.$$

Since $\{s_n\}$ is an increasing sequence,

$$\sum_{k=n+1}^{n+p} \frac{a_k}{s_k} \geq \frac{\sum_{k=n+1}^{n+p} a_k}{s_{n+p}} = \frac{s_{n+p} - s_n}{s_{n+p}} = 1 - \frac{s_n}{s_{n+p}}. \tag{6.3}$$

But $s_{n+p} \to \infty$ as $p \to \infty$. Thus, p may be chosen so large that $s_{n+p} > 2s_n$. For such a choice of p, it follows from (6.3) that

$$\sum_{k=n+1}^{n+p} \frac{a_k}{s_k} > \frac{1}{2}.$$

Hence, $\sum_{n=1}^{\infty}(a_n/s_n)$ diverges. To prove (ii), observe that

$$\frac{a_n}{s_n^2} \leq \frac{a_n}{s_n s_{n-1}} = \frac{s_n - s_{n-1}}{s_n s_{n-1}} = \frac{1}{s_{n-1}} - \frac{1}{s_n}.$$

Applying the Cauchy criterion, we have

$$\sum_{k=n+1}^{n+p} \frac{a_k}{s_k^2} \leq \sum_{k=n+1}^{n+p} \left(\frac{1}{s_{k-1}} - \frac{1}{s_k}\right) = \frac{1}{s_n} - \frac{1}{s_{n+p}} < \frac{1}{s_n}.$$

For any preassigned $\epsilon > 0$, we may choose n large enough so that $1/s_n < \epsilon$. Therefore, $\sum_{n=1}^{\infty}(a_n/s_n^2)$ converges. ∎

Corollary 6.6. *The series* $\sum_{n=1}^{\infty}(1/n)$ *diverges and* $\sum_{n=1}^{\infty}(1/n^2)$ *converges.*

Proof. Apply Theorem 6.5, with $a_n \equiv 1$. ∎

Moreover, a standard argument from real variable theory gives that

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$$

converges for $\alpha > 1$.

Corollary 6.7. *If* $a_n > 0$ *and* $\sum_{n=1}^{\infty} a_n$ *diverges, then there exists a positive sequence* $\{b_n\}$ *such that* $b_n/a_n \to 0$ *and* $\sum_{n=1}^{\infty} b_n$ *diverges.*

Proof. Choose $b_n = a_n/s_n$, and apply Theorem 6.5. ∎

This corollary shows that there is no "slowest" diverging series. The reader will be asked, in Exercise 6.19(4) to show that there is no slowest converging series.

We shall return to series after establishing additional properties for sequences. We prove the obvious, but useful

Theorem 6.8. (Mousetrap Principle) *Let* $0 \le a_n \le b_n$ *for every* n, *and assume that* $\lim_{n\to\infty} b_n = 0$. *Then* $\lim_{n\to\infty} a_n = 0$.

Proof. Given $\epsilon > 0$, there exists an integer N such that $|b_n| < \epsilon$ for $n > N$. But then we also have $|a_n| \le |b_n| < \epsilon$ for $n > N$. Hence, $a_n \to 0$. ∎

Theorem 6.9. *For* $\alpha > 0$, $\beta > 0$, *and* x *real, we have*

(i) $\displaystyle\lim_{x\to\infty} \frac{(\ln x)^\alpha}{x^\beta} = 0$;

(ii) $\displaystyle\lim_{x\to\infty} \frac{x^\alpha}{e^{\beta x}} = 0$.

Proof. Setting $f(x) = \ln x$ and $g(x) = x^{\beta/\alpha}$, we apply l'Hôpital's rule for functions of a real variable to obtain

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to\infty} \frac{1}{(\beta/\alpha)x^{\beta/\alpha}} = 0.$$

Therefore,

$$\left(\frac{\ln x}{x^{\beta/\alpha}}\right)^\alpha = \frac{(\ln x)^\alpha}{x^\beta} \to 0 \quad \text{as } x \to \infty,$$

and (i) is proved.

Letting $y = \ln x$ in (i), and observing that $y \to \infty$ as $x \to \infty$, we obtain (ii). ∎

Corollary 6.10. *For* $\alpha > 0$ *and* $\beta > 0$, *we have*

(i) $\displaystyle\lim_{n\to\infty} \frac{(\ln n)^\alpha}{n^\beta} = 0$;

(ii) $\displaystyle\lim_{n\to\infty} \frac{n^\alpha}{e^{\beta n}} = 0$;

(iii) $\displaystyle\lim_{n\to\infty} n^{1/n} = 1$;

(iv) $\displaystyle\lim_{n\to\infty} r^n = 0 \qquad (|r| < 1)$.

Proof. Letting $x = n$ in Theorem 6.9, we obtain (i) and (ii). Next, $n^{1/n} \to 1$ if $(\ln n)/n \to 0$. But this is a special case of (i), and (iii) is proved. Finally, setting $\beta = -\ln|r|$ in (ii), we have

$$\frac{n^\alpha}{e^{-n\ln|r|}} = n^\alpha|r|^n \ge |r|^n.$$

Finally, (iv) follows from (ii) and the mousetrap principle. ∎

Our next example is useful for establishing the convergence of many series. Consider the geometric series $\sum_{k=0}^{\infty} z^k$. Then we may express the partial sum $s_n = 1 + z + \cdots + z^{n-1}$ in explicit form. If $z = 1$, then $s_n = n$; if $z \neq 1$, then we have

$$s_n = \sum_{k=1}^{n} z^{k-1} = \frac{1-z^n}{1-z} = \frac{1}{1-z} - \frac{z^n}{1-z}.$$

When $z = 1$, $\{s_n\}$ is unbounded and so has no limit. Thus, the geometric series diverges for $z = 1$. For $z \neq 1$, we may use the convergence of $\{z^n\}$ to zero for $|z| < 1$ to show that

$$\sum_{k=1}^{\infty} z^{k-1} = \frac{1}{1-z} \quad (|z| < 1).$$

On the other hand, if $|z| \geq 1$, then the kth term z^{k-1} does not converge to 0 so that the series does not converge. To summarize

$$\lim_{n\to\infty} s_n = \sum_{k=0}^{\infty} z^k = \begin{cases} \dfrac{1}{1-z} & \text{for } |z| < 1 \\ \text{diverges} & \text{for } |z| \geq 1 \end{cases}$$

and the convergence of $\sum_{k=0}^{\infty} z^k$ is absolute when $|z| < 1$. Hence, a series $\sum_{n=1}^{\infty} a_n$ converges absolutely if there exists a constant r, $0 \leq r < 1$, and a real number M such that $|a_n| \leq Mr^n$ for $n > N$. For example, consider the series

$$\sum_{n=1}^{\infty} a_n, \quad a_n = \frac{1}{3^n} + i\frac{1}{4^n},$$

which converges. But then, how do we sum the series? In view of Proposition 6.2, we have

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} 3^{-n} + i\sum_{n=1}^{\infty} 4^{-n} = \frac{1/3}{1-1/3} + i\frac{1/4}{1-1/4} = \frac{1}{2} + \frac{i}{3}.$$

The convergence properties of complex series may be deduced from those of real series. If $\{a_k\}$ is a sequence of complex numbers, we write $a_k = \alpha_k + i\beta_k$, where $\{\alpha_k\}$ and $\{\beta_k\}$ are sequences of real numbers. By Theorem 2.12, we have that the complex series $\sum_{k=1}^{\infty} a_k$ converges if and only if $\sum_{k=1}^{\infty} \alpha_k$ and $\sum_{k=1}^{\infty} \beta_k$ both converge. That is,

$$\sum_{k=1}^{\infty} (\alpha_k + i\beta_k) = \alpha + i\beta \iff \sum_{k=1}^{\infty} \alpha_k = \alpha \text{ and } \sum_{k=1}^{\infty} \beta_k = \beta.$$

For instance, the series

$$\sum_{k=1}^{\infty} \frac{1 + i\cos(1/k)}{3^k}$$

converges.

If, in the series $\sum a_k$, the summation range is $(-\infty, \infty)$ instead of $(0, \infty)$ (or $(1, \infty)$ or (p, ∞), where p is a fixed integer), then we have a series of complex numbers a_k, namely, $\sum_{k=-\infty}^{\infty} a_k$. The most efficient way of handling this is to discuss separately the two series

$$\sum_{k=0}^{\infty} a_k \quad \text{and} \quad \sum_{k=-\infty}^{-1} a_k = \sum_{k=1}^{\infty} a_{-k}$$

so that the convergence of this series depends on the convergence of both series. This approach helps us to formulate the following definition. *"A series $\sum_{k=-\infty}^{\infty} a_k$ converges if and only if both $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=1}^{\infty} a_{-k}$ converge."* In other words, we write

$$\sum_{k=-\infty}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=0}^{n} a_k + \lim_{m\to\infty} \sum_{k=1}^{m} a_{-k}$$

provided both the limits on the right exist. Note that m and n tend independently to ∞. According to the definition, it is clear that if $\sum_{k=-\infty}^{\infty} |a_k|$ converges, then $\sum_{k=-\infty}^{\infty} a_k$ converges. Moreover, $\sum_{k=-\infty}^{\infty} |a_k|$ converges precisely when both $\sum_{k=0}^{\infty} |a_k|$ and $\sum_{k=1}^{\infty} |a_{-k}|$ converge.

Suppose $\{a_n\}$ is a bounded sequence of complex numbers, and that A is the set of subsequential limits of $\{a_n\}$. Some properties of A have already been discussed. For instance, the set A is nonempty (Theorem 2.17) and consists of one point if and only if $\{a_n\}$ converges (Theorem 2.14). To insure a pleasant treatment of Taylor and Laurent series, it is necessary to introduce the so-called "limit superior" of a sequence $\{a_n\}$ of nonnegative real numbers. If the sequence $\{a_n\}$ is real and bounded, the set A has a least upper bound (Dedekind property).

Let $\{a_n\}$ be a real bounded sequence, and let A be the set of subsequential limits of $\{a_n\}$. Setting $a^* = \operatorname{lub} A$, we call a^* the *limit superior* of $\{a_n\}$, and write

$$\limsup_{n\to\infty} a_n = a^* \quad \text{or} \quad \overline{\lim}_{n\to\infty} a_n = a^*.$$

This definition may be extended to unbounded real sequences. If $\{a_n\}$ is unbounded above, we say that

$$\limsup_{n\to\infty} a_n = \infty;$$

whereas, if all but a finite number of a_n are less than any preassigned real number, we say that

$$\limsup_{n\to\infty} a_n = -\infty.$$

A useful counterpart to the limit superior is the *limit inferior*. For a bounded real sequence $\{a_n\}$ we set a_*=glb{ subsequential limits }, and write

$$\liminf_{n\to\infty} a_n = a_* \quad \text{or} \quad \underline{\lim}_{n\to\infty} a_n = a_*.$$

If $\{a_n\}$ is unbounded below, we say that

$$\liminf_{n\to\infty} a_n = -\infty;$$

whereas, if all but a finite number of a_n are greater than any preassigned real number, we say that

$$\liminf_{n\to\infty} a_n = \infty.$$

Formulating theorems in terms of limit superior or limit inferior, rather than in terms of limit, has one distinct advantage. In the extended real number system ($\pm\infty$ included), the limit superior and limit inferior of a real sequence always exist. This enables us to prove theorems about sequences without worrying about the existence of limits.

Examples 6.11. (i) If $a_n = \frac{1}{n}$, then $\limsup_{n\to\infty} a_n = 0$.

(ii) If $a_n = (-1)^n$, then $\limsup_{n\to\infty} a_n = 1$ and $\liminf_{n\to\infty} a_n = -1$.

(iii) If $a_n = 3^n$, then $\{a_n\}$ is not bounded above and $\limsup_{n\to\infty} a_n = \infty$.

(iv) If $a_n = -n+(-1)^n n$, then the sequence $-2, 0, -2(3), 0, -2(5), \dots$ is not bounded below. Thus, $\limsup_{n\to\infty} a_n = 0$ and $\liminf_{n\to\infty} a_n = -\infty$.

(v) If $a_n = 1 - (1/2)^n$ for $n \in \mathbb{N}$, then $\limsup_{n\to\infty} a_n = 1$.

(vi) If $a_n = (1+c)^n$ with $c > 0$ and $n \in \mathbb{N}$, then $\limsup_{n\to\infty} a_n = \infty$.

(vii) Let $\{a_n\} = n\sin^2(n\pi/4)$. Since $0 \le a_n < \infty$ for every n, no subsequence of $\{a_n\}$ can approach a value less than 0. In order to show that

$$\limsup_{n\to\infty} a_n = \infty \quad \text{and} \quad \liminf_{n\to\infty} a_n = 0,$$

it suffices to find one (out of many) subsequences that approach the desired value. We have

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} a_{4n+2} = \lim_{n\to\infty} (4n+2) = \infty$$

and

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_{4n} = \lim_{n\to\infty} 4n.0 = 0.$$

(viii) Let $a_n = (1 + 1/n)\cos n\pi$. Then

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} a_{2n} = 1 \quad \text{and} \quad \liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_{2n+1} = -1.$$

(ix) Let $a_n = \sin(n\pi/2) + \sin(n\pi/4)$. Then

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} a_{4n+1} = 1 + \frac{\sqrt{2}}{2}$$

and

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_{4n-1} = -1 - \frac{\sqrt{2}}{2}.$$

(x) Let $a_n = n\cos n\pi$. Then

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} a_{2n} = \infty \quad\text{and}\quad \liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_{2n-1} = -\infty.$$

(xi) Let $a_n = 5n\cos n\pi - n^2$. Then $a_n \le 5n - n^2$ for every n, so that

$$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = -\infty. \qquad \bullet$$

It is clear that the limit superior and limit inferior are both unique, and that a real sequence $\{a_n\}$ converges to L, L finite, if and only if

$$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = L.$$

But some of the standard limit theorems are false for both the limit superior and the limit inferior. For instance, if $a_n = (-1)^n$ and $b_n = (-1)^{n+1}$, then

$$\limsup_{n\to\infty}(a_n + b_n) = 0 \ne \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n = 2$$

and

$$\liminf_{n\to\infty}(a_n + b_n) = 0 \ne \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n = -2.$$

However, we do have the following inequalities:

Theorem 6.12. *Let $\{a_n\}$ and $\{b_n\}$ be real, bounded sequences. Then, we have*

(i) $\limsup_{n\to\infty}(a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$

(ii) $\liminf_{n\to\infty}(a_n + b_n) \ge \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$.

Proof. Let $\limsup_{n\to\infty} a_n = a$ and $\limsup_{n\to\infty} b_n = b$. Assume inequality (i) is false. Then for some $\epsilon > 0$ there is a subsequence $\{a_{n_k} + b_{n_k}\}$ of $\{a_n + b_n\}$ such that $a_{n_k} + b_{n_k} > a + b + \epsilon$ for all n_k. But then either

$$a_{n_k} > a + \frac{\epsilon}{2} \quad\text{or}\quad b_{n_k} > b + \frac{\epsilon}{2}$$

infinitely often. This implies that either

$$\limsup_{n\to\infty} a_n \ge a + \frac{\epsilon}{2} \quad\text{or}\quad \limsup_{n\to\infty} b_n \ge b + \frac{\epsilon}{2}$$

(Why?). This contradicts our assumption, and (i) is proved. The proof of (ii) is similar, and will be omitted. ∎

It turns out that absolute convergence for (power) series plays a central role in complex analysis as it is much easier to test for absolute convergence than for convergence by other means. The nth root test is one such useful result in determining convergence properties of power series.

Theorem 6.13. (Root test) *Let $\{a_n\}_{n\geq 1}$ be a complex sequence, and suppose that*

$$\limsup_{n\to\infty} |a_n|^{1/n} = L.$$

Then $\sum_{n=1}^{\infty} a_n$ converges absolutely if $L < 1$, and diverges if $L > 1$.

Proof. If $L < 1$, choose r such that $L < r < 1$. Then, we have $|a_n|^{1/n} < r$ for large values of n; that is,

$$|a_n| < r^n \quad \text{for } n > N.$$

The convergence of $\sum_{n=1}^{\infty} |a_n|$ now follows from the convergence of $\sum_{n=1}^{\infty} r^n$ ($|r| < 1$).

If $L > 1$, then $|a_n|^{1/n} > 1$ for infinitely many values of n. But then $|a_n| > 1$ infinitely often. Hence, $a_n \not\to 0$ and the series $\sum_{n=1}^{\infty} a_n$ cannot converge. ∎

Remark 6.14. When $L = 1$, the root test gives no information. The series $\sum_{n=1}^{\infty}(1/n)$ diverges and $\sum_{n=1}^{\infty}(1/n^2)$ converges. However,

$$\limsup_{n\to\infty} \left|\frac{1}{n}\right|^{1/n} = \limsup_{n\to\infty} \left|\frac{1}{n^2}\right|^{1/n} = 1. \quad \bullet$$

Theorem 6.15. (Ratio test) *Let $\{a_n\}_{n\geq 1}$ be a complex sequence, and suppose that*

$$\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = L \quad \textit{and} \quad \liminf_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = l.$$

Then $\sum_{n=1}^{\infty} a_n$ converges absolutely if $L < 1$, and diverges if $l > 1$. The test offers no conclusion concerning the convergence of the series if $l \leq 1 \leq L$.

We leave the proof as an exercise. To show that the test fails in the last case, we simply consider the series in Remark 6.14 and observe that $L = l = 1$, in both cases.

As usual we easily have the following corollary.

Corollary 6.16. *Let $\{a_n\}_{n\geq 1}$ be a complex sequence, and suppose that*

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = L. \tag{6.4}$$

Then $\sum_{n=1}^{\infty} a_n$ converges absolutely if $L < 1$, and diverges if $L > 1$. The test offers no conclusion concerning the convergence of the series if $L = 1$.

Clearly, from Theorem 6.13, the conclusion of the corollary continues to hold if (6.4) is replaced by $\lim_{n\to\infty} |a_n|^{1/n} = L$.

Example 6.17. Consider the series $\sum_{n=1}^{\infty}(1/n!)$. Then the ratio test is easier to apply and it is not obvious that $\limsup_{n\to\infty} |(1/n!)|^{1/n} < 1$. Similarly, the ratio test may be easier to examine the convergence property of the series $\sum_{n=1}^{\infty}(-1)^n(1-i)^n/n!$. •

Questions 6.18.

1. What types of theorems for real sequences remain valid for complex sequences?
2. Does there exist a convergent series $\sum_{n=1}^{\infty} a_n$ for which $\lim_{n\to\infty} a_n \neq 0$?
3. Does there exist a divergent series $\sum_{n=1}^{\infty} a_n$ for which $\lim_{n\to\infty} a_n = 0$?
4. Do the series $\sum_{n=N}^{\infty} a_n$ and $\sum_{n=1}^{\infty} a_n$ converge or diverge together?
5. What does it mean to say that there is no slowest converging series?
6. Why is the Mousetrap Principle so named?
7. For which convergent series can you determine the sum?
8. What alternate definitions for limit superior and limit inferior might we have given?
9. Can limit inferior be defined in terms of limit superior?
10. What are some advantages and disadvantages of allowing the limit superior of a sequence to assume the values $\pm\infty$?
11. How does the limit superior of a product compare with the product of limit superiors?
12. Can Theorem 6.12 be modified to include unbounded sequences?
13. If $|a_n|^{1/n} < 1$ for every $n \geq 1$, does $\sum_{n=1}^{\infty} a_n$ necessarily converge?
14. If $z_n \to z_0$ and $w_n \to w_0$, does $\frac{1}{n}\sum_{k=1}^{n} z_k w_k \to z_0 w_0$?
15. If $z_n \to z_0$, does $\frac{1}{2^n}\sum_{k=1}^{n} \binom{n}{k} z_k \to z_0$?

Exercises 6.19.

1. Let $a_n = \alpha_n + i\beta_n$, α_n and β_n real. Show that $\sum_{n=1}^{\infty} |a_n|$ converges if and only if both $\sum_{n=1}^{\infty} |\alpha_n|$ converges and $\sum_{n=1}^{\infty} |\beta_n|$ converges.
2. Let $s_n = \sum_{k=1}^{n} a_k$. If $a_k = 1/k$, show that $s_{2^{n+1}} - s_{2^n} > \frac{1}{2}$ for every n. If $a_k = (-1)^{k+1}/k$, show that $|s_{n+p} - s_n| < 2/n$ for every n and p.
3. Suppose $a_n > 0$ for every n. Show that $\sum_{n=1}^{\infty} a_n$ diverges if and only if for any integers M and N, there exists an integer p such that $\sum_{n=N}^{N+p} a_n > M$.
4. Let $a_n > 0$, and suppose $\sum_{n=1}^{\infty} a_n$ converges. If $r_n = \sum_{k=n}^{\infty} a_k$, show that

$$\text{(a)} \quad \sum_{n=1}^{\infty} \frac{a_n}{r_n} \quad \text{diverges} \qquad\qquad \text{(b)} \quad \sum_{n=1}^{\infty} \frac{a_n}{\sqrt{r_n}} \quad \text{converges.}$$

5. Let A be the set of subsequential limits of a complex sequence. Show that A is closed set.
6. Find the infimum and supremum of the following

$$\text{(i) } 5 + \sin(n\pi/3) \qquad \text{(ii) } 1/n + \sin(n\pi/3) \qquad \text{(iii) } 1/n + \cos(n\pi/3).$$

7. Suppose that $\{a_n\}$ is a real sequence and $\lim_{n\to\infty} a_n = a$, $a \neq 0$. For any sequence $\{b_n\}$, show that
(a) $\limsup_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$

(b) $\liminf_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$
(c) $\limsup_{n\to\infty} a_n b_n = \lim_{n\to\infty} a_n \limsup_{n\to\infty} b_n$
(d) $\liminf_{n\to\infty} a_n b_n = \lim_{n\to\infty} a_n \liminf_{n\to\infty} b_n$.

8. Show that $\limsup_{n\to\infty} a_n = L$, L finite, if and only if the following conditions hold: For any $\epsilon > 0$,
   a) $a_n < L + \epsilon$ for all but a finite number of n;
   b) $a_n > L - \epsilon$ infinitely often.
9. Show that $\liminf_{n\to\infty} a_n = L$, L finite, if and only if the following conditions hold: For any $\epsilon > 0$,
   a) $a_n < L + \epsilon$ infinitely often;
   b) $a_n > L - \epsilon$ for all but a finite number of n.
10. Let $\{a_n\}$ be a complex sequence.
    a) If $\limsup_{n\to\infty} |a_{n+1}/a_n| = L < 1$, show that $\sum_{n=1}^{\infty} a_n$ converges absolutely.
    b) $\liminf_{n\to\infty} |a_{n+1}/a_n| = L > 1$, show that $\sum_{n=1}^{\infty} a_n$ diverges.

    Show by an example that this result provides no information about the convergence or divergence of the series when $L = 1$.
11. Suppose $a_n > 0$ for every n. Show that

$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} \le \liminf_{n\to\infty} a_n^{1/n} \le \limsup_{n\to\infty} a_n^{1/n} \le \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

6.2 Uniform Convergence

A sequence of functions $\{f_n\}$ *converges pointwise* to a function f on a set E ($f_n \to f$) if to each $z_0 \in E$ and $\epsilon > 0$, there corresponds an integer $N = N(\epsilon, z_0)$ for which

$$|f_n(z) - f(z_0)| < \epsilon \quad \text{whenever} \quad n > N(\epsilon, z_0).$$

To say that a sequence of functions $\{f_n\}$ converges pointwise on a set E is equivalent to saying that the sequence of numbers $\{f_n(z_0)\}$ converges for each $z_0 \in E$. The limit function f is then defined by

$$\lim_{n\to\infty} f_n(z_0) = f(z_0) \quad (z_0 \in E).$$

The integer N in the definition of pointwise convergence may, in general, vary with the points in the set. If, however, one integer can be found that works for all such points, the convergence is said to be *uniform*. That is, a sequence of functions $\{f_n\}$ *converges uniformly* to f on a set E ($f_n \rightrightarrows f$) if to each $\epsilon > 0$, there corresponds an integer $N = N(\epsilon)$ such that, for all $z \in E$,

$$|f_n(z) - f(z)| < \epsilon \quad \text{whenever } n > N(\epsilon).$$

Thus, when n is large, $f_n(z)$ is required to be uniformly "close" to $f(z)$ on E.

We emphasize that uniform convergence on a set implies (pointwise) convergence. A formulation of the negation of uniform convergence will be helpful when producing examples that show the converse to be false. The convergence of $\{f_n\}$ to f on E is *not* uniform if there exists an $\epsilon > 0$ such that to each integer N there corresponds an integer $n(> N)$ and a point $z_n \in E$ for which $|f_n(z_n) - f(z_n)| \geq \epsilon$.

Recall the distinction between continuity and uniform continuity. A continuous function is uniformly continuous on a set if a single $\delta = \delta(\epsilon)$ can be found that works for all points in the set. In Chapter 2 the function $f(z) = 1/z$ was shown to be continuous, but not uniformly continuous, on the set $0 < |z| < 1$. The following example is an analog for convergence.

Example 6.20. Let $f_n(z) = 1/(nz)$. Then we see that the sequence $\{f_n(z)\}$ converges pointwise, but not uniformly, to the function $f(z) \equiv 0$ on the set $0 < |z| < 1$.

For a given $\epsilon > 0$, in order that

$$|f_n(z) - 0| = |1/(nz)| < \epsilon,$$

it is necessary that $n > 1/(\epsilon|z|)$. So the corresponding $N = N(z;\epsilon)$ is an integer greater than $1/(\epsilon|z|)$. Note that, as $|z|$ decreases, the corresponding N increases without bound. Thus, we say that the sequence converge pointwise but not uniformly on $\{z : 0 < |z| < 1\}$.

Alternatively, we argue in the following manner. If this convergence were uniform, there would exist an integer N for which the inequality $|1/Nz| < \epsilon < 1$ would be valid for all z, $0 < |z| < 1$. But the inequality does not hold for $z = 1/N$. ●

We have shown that the convergence in the above example is not uniform because

$$\left| f_n\left(\frac{1}{n}\right) - f\left(\frac{1}{n}\right) \right| = 1 \quad \text{for all} \quad n.$$

Example 6.21. Let $f_n(z) = 1/(1+nz)$. Then the sequence $\{f_n(z)\}$ converges uniformly in the region $|z| \geq 2$, but does not converge uniformly in the region $|z| \leq 2$. Indeed, the sequence $\{f_n\}$ converges pointwise everywhere to the function

$$f(z) = \begin{cases} 0 & \text{if } z \neq 0 \\ 1 & \text{if } z = 0. \end{cases}$$

If $|z| \geq 2$, then

$$|f_n(z)| = \left| \frac{1}{1+nz} \right| \leq \frac{1}{|nz| - 1} \leq \frac{1}{2n-1} \leq \frac{1}{n}.$$

Therefore, $|f_n(z)| < \epsilon$ whenever $n > 1/\epsilon$, which proves uniform convergence in the region $|z| \geq 2$.

If the convergence were uniform for $|z| \le 2$, there would exist an integer N for which the inequality $|f_N(z) - f(z)| < \frac{1}{2}$ would be valid for all z in the region. But

$$\left| f_n\left(\frac{1}{n}\right) - f\left(\frac{1}{n}\right) \right| = \left| \frac{1}{1 + n \cdot (1/n)} - 0 \right| = \frac{1}{2} \quad \text{for every } n. \qquad \bullet$$

Remark 6.22. Theorem 2.44 states that a function continuous on a compact set is uniformly continuous. The above example shows that pointwise convergence, even on a compact set, need not imply uniform convergence. ●

Example 6.23. Set $f_n(z) = z^n$. Then the sequence $\{f_n(z)\}$ converges pointwise on the set $|z| < 1$ and uniformly on the set $|z| \le r < 1$.

The pointwise convergence for $|z| < 1$ follows from Corollary 6.10(iv). Note that for $r < 1$, given $\epsilon > 0$,

$$r^n < \epsilon \iff n > \frac{\ln \epsilon}{\ln r}.$$

Since $r^n \to 0$ $(r < 1)$, an integer $N = N(\epsilon)$ may be found for which $r^n < \epsilon$ $(n > N > \frac{\ln \epsilon}{\ln r})$. But then,

$$|z^n| \le r^n < \epsilon \quad (|z| \le r, n > N(\epsilon)).$$

Hence, $\{f_n(z)\}$ converges uniformly to zero in the disk $|z| \le r$. Remember that the choice of N with $N > \frac{\ln \epsilon}{\ln r}$ is possible for an arbitrary $\epsilon > 0$ and $0 < r < 1$. However, for each fixed $\epsilon > 0$, as $r \to 1^-$, N must be increased without bound. It follows that the convergence is not uniform for $|z| < 1$.

Alternatively, if the convergence were uniform on the set $|z| < 1$, then for sufficiently large n the inequality $|z^n| < \epsilon$ would be valid for all z, $|z| < 1$. Choosing $z = 1 - 1/n$, we have $z^n = (1 - 1/n)^n$, and, from elementary calculus

$$\left(1 - \frac{1}{n}\right)^n \to \frac{1}{e} \quad \text{as} \quad n \to \infty.$$

Hence, $|z^n| > \frac{1}{3}$ $(z = 1 - 1/n, n > N)$, and the convergence is not uniform on $|z| < 1$. ●

Example 6.24. Let $f_n(z) = z/n$ on $\mathbb{C}$. Then $f_n(z) \to f(z) \equiv 0$ on $\mathbb{C}$ but not uniformly. Note that for a given $\epsilon > 0$

$$|f_n(z) - 0| = \left| \frac{z}{n} \right| < \epsilon \iff n > \frac{|z|}{\epsilon}$$

showing that, as $|z|$ increases, the corresponding N increases without bound. It follows that the sequence converges pointwise but not uniformly on $\mathbb{C}$. On the other hand if we restrict the domain to a bounded subset of $\mathbb{C}$, say $\Omega = \{z \in \mathbb{C} : |z| < 2006\}$, then the convergence is uniform. ●

Example 6.25. We wish to show the sequence $f_n(z) = z/n^2$ converges uniformly to $f(z) \equiv 0$ for $|z| \le R$, but does not converge uniformly in the plane.

To do this, we recall that for $|z| \le R$, the inequality

$$|f_n(z)| \le \frac{R}{n^2} < \epsilon \quad (n > \sqrt{R/\epsilon})$$

shows the convergence to be uniform on any disk $|z| \le R$. But if the convergence were uniform in $\mathbb{C}$, then we would have $|z/N^2| < 1$ for some integer N and for all z. Choosing $z = N^2$ elicits the appropriate contradiction. ●

The uniform convergence of a sequence $\{f_n\}$ on E is often deduced from the convergence of $\{f_n\}$ at some point $z = z_0$ by showing that the inequality

$$|f_n(z) - f(z)| \le |f(z_0) - f(z)|$$

is valid for all $z \in E$.

If a sequence does not converge uniformly, there is usually some "bad point" to be exploited. In Examples 6.20, 6.21, 6.23, and 6.25 the bad points were, respectively, $z = 0$, $z = 0$, $z = 1$, and $z = \infty$. For each such point z_0, the expression $\lim_{z\to z_0}(\lim_{n\to\infty} f_n(z))$ could not be replaced by

$$\lim_{n\to\infty}(\lim_{z\to z_0} f_n(z)).$$

In Example 6.21, for instance,

$$\lim_{z\to 0}\left(\lim_{n\to\infty}\frac{1}{1+nz}\right) = 0 \quad\text{while}\quad \lim_{n\to\infty}\left(\lim_{z\to 0}\frac{1}{1+nz}\right) = 1.$$

The importance of uniform convergence is that it does allow for the interchange of many limit operations. This, in turn, compels the limit function to retain many properties of the sequence.

Theorem 6.26. *Suppose $\{f_n\}$ converges uniformly to a function f on E. If each f_n is continuous at a point $z_0 \in E$, then the limit function f is also continuous at z_0. That is,*

$$\lim_{z\to z_0}\left(\lim_{n\to\infty} f_n(z)\right) = \lim_{n\to\infty}\left(\lim_{z\to z_0} f_n(z)\right).$$

Proof. We must show that for any $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(z) - f(z_0)| < \epsilon$ for all z in E satisfying $|z - z_0| < \delta$. The inequality

$$|f(z) - f(z_0)| \le |f(z) - f_n(z)| + |f_n(z) - f_n(z_0)| + |f_n(z_0) - f(z_0)| \quad (6.5)$$

is valid for every n. The uniform convergence of $\{f_n\}$ allows us to choose N independent of z so that

$$|f(z) - f_N(z)| < \frac{\epsilon}{3} \quad (z \in E).$$

Letting $n = N$ in (6.5), we have

$$|f(z) - f(z_0)| < \frac{\epsilon}{3} + |f_N(z) - f_N(z_0)| + \frac{\epsilon}{3}. \tag{6.6}$$

By the continuity of f_N at $z = z_0$,

$$|f_N(z) - f_N(z_0)| < \frac{\epsilon}{3} \tag{6.7}$$

for z sufficiently close to z_0. Combining (6.6) and (6.7), the desired result follows. ■

Remark 6.27. Theorem 6.26 furnishes us with a necessary, but not sufficient, condition for uniform convergence. In Example 6.23, it was shown that the sequence of continuous functions $\{z^n\}$ converges nonuniformly in the region $|z| < 1$ to the continuous function $f(z) \equiv 0$. However, in Example 6.21, the discontinuity of the limit function at $z = 0$ rules out uniform convergence for the sequence $\{1/(1 + nz)\}$ in any region containing the origin. ●

Our definition and discussion of uniform convergence remains valid for real-valued functions of a real variable. In fact, the same conclusions may be drawn from the preceding four examples when z is replaced by x and the regions are replaced by their corresponding intervals. Moreover, there is an interesting geometric interpretation to uniform convergence of real-valued functions. If $\{f_n(x)\}$ converges uniformly to $f(x)$ on a set E, then for sufficiently large n we have

$$f(x) - \epsilon < f_n(x) < f(x) + \epsilon \quad \text{for all } x \text{ in } E.$$

This means that there is some curvilinear strip of vertical width 2ϵ that contains the graph of all functions $y = f_n(x)$, with $n > N$, and that each such curve is never a distance more than ϵ away from the curve $y = f(x)$.

Example 6.28. Let $f_n(x) = x^2 + \sin nx/n$. We have

$$|f_n(x) - x^2| = \left|\frac{\sin nx}{n}\right| \le \frac{1}{n} \quad (x \text{ real}),$$

and $|f_n(x) - x^2| < \epsilon$ for $n > 1/\epsilon$. Hence, the sequence $\{f_n(x)\}$ converges uniformly to $f(x) = x^2$ on the set of real numbers (see Fig. 6.1). ●

Certainly no discussion of convergence is complete without a Cauchy criterion. Rewording Theorem 2.20 for functions, we have that the sequence $\{f_n(z)\}$ convergence pointwise on E if and only if $\{f_n(z_0)\}$ is a Cauchy sequence for each $z_0 \in E$. That is, to each $z_0 \in E$ and $\epsilon > 0$, there corresponds an integer $N = N(\epsilon, z_0)$ for which $|f_n(z_0) - f_m(z_0)| < \epsilon$ whenever $n, m > N(\epsilon, z_0)$.

A sequence $\{f_n\}$ is said to be *uniformly Cauchy* on E if to each $\epsilon > 0$ there corresponds an integer $N = N(\epsilon)$ such that, for all $z \in E$,

$$|f_n(z) - f_m(z)| < \epsilon \quad \text{whenever} \quad n, m > N(\epsilon).$$

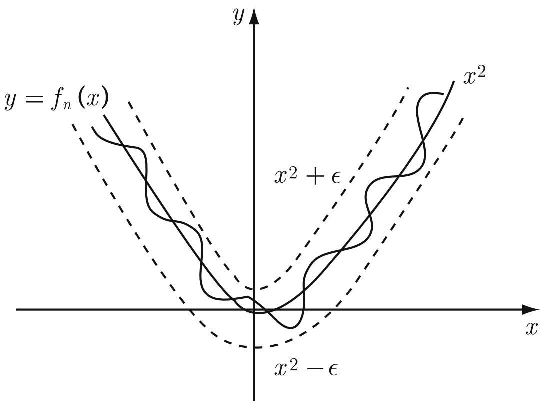


Figure 6.1. Illustration for uniform convergence

Theorem 6.29. *A sequence of functions converges uniformly on a set E if and only if the sequence is uniformly Cauchy on E.*

Proof. Suppose $\{f_n\}$ converges uniformly to f on E. Then, given $\epsilon > 0$,

$$|f_n(z) - f(z)| < \frac{\epsilon}{2} \quad \text{for} \quad n > N(\epsilon) \quad \text{and for all} \quad z \in E.$$

But then, for $n, m > N(\epsilon)$, we have

$$|f_n(z) - f_m(z)| \leq |f_n(z) - f(z)| + |f_m(z) - f(z)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence, the sequence $\{f_n\}$ is uniformly Cauchy on E.

Conversely, suppose $\{f_n\}$ is uniformly Cauchy on E. In particular, $\{f_n(z_0)\}$ is a Cauchy sequence for each $z_0 \in E$, and thus $\{f_n\}$ converges pointwise to a function f. We wish to show that this convergence is uniform. Given $\epsilon > 0$, there exists an integer $N = N(\epsilon)$ such that $n, m > N$ implies

$$|f_n(z) - f_m(z)| < \frac{\epsilon}{2} \quad \text{for all} \quad z \in E. \tag{6.8}$$

Fixing $n(> N)$ and letting m vary, (6.8) leads to

$$|f_n(z) - f(z)| = \lim_{m\to\infty} |f_n(z) - f_m(z)| \leq \frac{\epsilon}{2} < \epsilon. \tag{6.9}$$

Since (6.9) is valid for all $z \in E$ and $n > N(\epsilon)$, the convergence of $\{f_n\}$ to f is uniform on E. ∎

In the previous section, we saw that properties for sequences of complex numbers could be reworded as properties for series of complex numbers. The remainder of this section will be devoted to the conversion from sequences of complex functions to series of complex functions. As expected, our work will parallel that of the previous section.

Given a sequence of functions $\{f_n(z)\}$ defined on a set E, we associate a new sequence of functions $\{S_n(z)\}$ defined by

$$S_n(z) = \sum_{k=1}^{n} f_k(z). \tag{6.10}$$

For all values of z for which $\lim_{n\to\infty} S_n(z)$ exists, we say that the series, denoted by $\sum_{k=1}^{\infty} f_k(z)$, converges, and write

$$f(z) = \lim_{n\to\infty} S_n(z) = \sum_{k=1}^{\infty} f_k(z).$$

If $\{S_n(z)\}$ converges uniformly on a set E, then the series $\sum_{k=1}^{\infty} f_k(z)$ is said to be uniformly convergent on E. Further, the series $\sum_{k=1}^{\infty} f_k(z)$ is absolutely convergent if $\sum_{k=1}^{\infty} |f_k(z)|$ converges. Moreover, a necessary condition for $\sum_{k=1}^{\infty} f_k(z)$ to converge uniformly on E is that $f_k(z) \to 0$ on E. This fact is evident if we write

$$f_n = S_n - S_{n-1}$$

and allow $n \to \infty$. Also, rewording the Cauchy criterion, we have

Theorem 6.30. *The series $\sum_{n=1}^{\infty} f_n(z)$ converges uniformly on a set E if and only if, to each $\epsilon > 0$, there corresponds an integer $N = N(\epsilon)$ such that for all $z \in E$, we have*

$$\left| \sum_{k=n+1}^{n+p} f_k(z) \right| < \epsilon \quad \textit{whenever} \quad n > N(\epsilon) \quad (p = 1, 2, 3, \dots).$$

Proof. Define S_n by (6.10), and apply Theorem 6.29. ∎

Theorem 6.30 may be used to establish a sufficient condition for the uniform convergence of a series, called the *Weierstrass M-test or dominated convergence test.*

Theorem 6.31. *Let $\{M_n\}_{n\geq 1}$ be a sequence of real numbers, and suppose that $|f_n(z)| \leq M_n$ for all $z \in E$ and each $n \in \mathbb{N}$. If $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n(z)$ converges uniformly (and absolutely) on the set E.*

Proof. That $\sum_{n=1}^{\infty} |f_n(z)|$ converges on E follows immediately from the comparison test for real series. To verify the uniform convergence of $\sum_{n=1}^{\infty} f_n(z)$ on E, we invoke the Cauchy criterion for $\sum_{n=1}^{\infty} M_n$. Thus, given $\epsilon > 0$, there exists an integer N such that, for $n > N$, we have

$$\sum_{k=n+1}^{n+p} M_k < \epsilon \quad (p = 1, 2, 3, \dots).$$

But

$$\left|\sum_{k=n+1}^{n+p} f_k(z)\right| \leq \sum_{k=n+1}^{n+p} |f_k(z)| \leq \sum_{k=n+1}^{n+p} M_k < \epsilon.$$

The uniform convergence now follows from Theorem 6.30. ■

Example 6.32. The geometric series $\sum_{n=0}^{\infty} z^n$ converges absolutely for $|z| < 1$ and uniformly for $|z| \leq r < 1$. Here we are dealing with a series of functions $\sum_{n=1}^{\infty} f_n(z)$ in the unit disk $|z| < 1$, where $f_n(z) = z^n$. We have

$$\left|\frac{1}{1-z}\right| = \left|\sum_{n=0}^{\infty} z^n\right| \leq \sum_{n=0}^{\infty} |z|^n = \frac{1}{1-|z|} \quad (|z| < 1),$$

and this proves absolutely convergence. Uniform convergence for $|z| \leq r$ then follows from the M-test (Theorem 6.31). Indeed, if we fix $r < 1$ and define $M_k = r^k$, then $\sum_{k=0}^{\infty} M_k$ converges, and $|z^k| \leq M_k$ for $|z| \leq r$. By the Weierstrass M-test, $\sum_{k=0}^{\infty} z^k$ converges uniformly for $|z| \leq r$, for each $r < 1$. We can also show that the series does not converge uniformly in the unit disk $|z| < 1$. Setting

$$S_n(z) = \sum_{k=0}^{n-1} z^k = \frac{1-z^n}{1-z},$$

the sequence $\{S_n(z)\}$ converges pointwise to $f(z) = 1/(1-z)$ for $|z| < 1$. Choosing $z = 1 - 1/n$, we have

$$|S_n(z) - f(z)| = \left|\frac{z^n}{1-z}\right| = n\left(1 - \frac{1}{n}\right)^n \to \infty \quad \text{as} \quad n \to \infty$$

(because $(1 - 1/n)^n \to 1/e$) showing that the partial sums do not converge uniformly for $|z| < 1$. Hence, the uniform convergence of the series cannot be extended to the disk $|z| < 1$. ●

Example 6.33. The series $\sum_{n=1}^{\infty} (\cos nz)/n^2$ converges uniformly and absolutely on the real line. Indeed, since $|\cos nz| \leq 1$ for all z real, Theorem 6.31 may be applied with $M_n = 1/n^2$ to obtain the desired result. By writing

$$\cos nz = (e^{inz} + e^{-inz})/2,$$

the reader may verify that $\cos nz/n^2$ does not approach 0 as n tends to ∞ unless z is real. Thus, $\sum_{n=1}^{\infty} (\cos nz)/n^2$ converges if and only if z is real. ●

Example 6.34. The series $\sum_{n=1}^{\infty} 2z^2/(n^2 + |z|)$ converges absolutely in the plane and uniformly for $|z| \leq R$, for each $R > 0$.

To see this, it suffices to observe that for any point z_0 in the plane,

$$\frac{2|z_0^2|}{n^2 + |z_0|} \leq \frac{2|z_0^2|}{n^2}.$$

Now, the absolute convergence in the plane is a consequence of the convergence of

$$2|z_0|^2 \sum_{n=1}^{\infty} \frac{1}{n^2},$$

and the uniform convergence on the disk $|z| \le R$ follows from the M-test, with $M_n = 2R^2/n^2$. ●

Example 6.35. We show that the Riemann zeta function

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

converges absolutely for $\operatorname{Re} z > 1$ and uniformly for $\operatorname{Re} z \ge 1+\epsilon$, $\epsilon > 0$. The given series is concerned with $\sum_{n=1}^{\infty} f_n(z)$, where $f_n(z) = n^{-z}$. We have

$$n^{-z} = e^{-z \operatorname{Log} n} = e^{-(x+iy) \operatorname{Log} n} \quad \text{and} \quad \left|n^{-z}\right| = e^{-x \ln n} = n^{-x}$$

so that $\sum_{n=1}^{\infty} |n^{-z}|$ converges for $x = \operatorname{Re} z > 1$. The uniform convergence for $\operatorname{Re} z \ge 1+\epsilon$ follows from the M-test, with $M_n = 1/n^{1+\epsilon}$. ●

Questions 6.36.

1. What kinds of sequences of functions converge uniformly in the plane?
2. Can a sequence of unbounded functions converge uniformly?
3. How would you define: $\{f_n\}$ converges uniformly to infinity?
4. Can a sequence of functions converge uniformly on every compact subset of a region and not converge uniformly in the region?
5. Can a sequence of functions converge pointwise on every compact subset of a region and not converge pointwise in the region?
6. Can a sequence of discontinuous functions converge uniformly to a continuous function?
7. Can a sequence of functions converge uniformly, but not absolutely, in a region? Absolutely, but not uniformly?
8. Suppose that for every $\epsilon > 0$, there exists an integer N such that $|f_N(z) - f(z)| < \epsilon$ for all z in E. Does $\{f_n\}$ converge uniformly to f in E? Does $\{f_n\}$ converge pointwise in E?
9. How would Theorem 6.31 be stated as a theorem for sequences?
10. If a sequence of differentiable functions converges uniformly, must the limit function be differentiable?

Exercises 6.37.

1. Show that $f_n = u_n + iv_n$ converges uniformly to $f = u + iv$ if and only if $\{u_n\}$ converges uniformly to u and $\{v_n\}$ converges uniformly to v.
2. Suppose $\{f_n\}$ converges uniformly to f and $\{g_n\}$ converges uniformly to g on E.

(a) Show that $\{f_n + g_n\}$ converges uniformly to $f + g$ on E.
(b) If, in addition, $|f_n| \le M$ and $|g_n| \le M$ for all $z \in E$ and all n, show that $\{f_n g_n\}$ converges uniformly to fg on E.

3. Show that $f_n(z) = \frac{z^n}{n}$ converges uniformly for $|z| < 1$. Show also that $f_n'(z)$ does not converge uniformly for $|z| < 1$ but it does converge uniformly for $|z| \le r$ for $r < 1$.
4. Suppose that $f(z)$ is unbounded on a set E. Let $f_n(z) \equiv f(z)$ for all n, and let $g_n(z) = 1/n$.
   a) Show that $f_n(z)$ and $g_n(z)$ both converge uniformly on the set E.
   b) Show that $f_n(z)g_n(z)$ converges pointwise, but not uniformly, on E.
5. Show that $\{f_n\}$ converges uniformly on a finite set if and only if $\{f_n\}$ converges pointwise.
6. Prove the following generalization of Theorem 6.26: Suppose $\{f_n\}$ converges uniformly to a function f on E, and f_n is continuous at $z_0 \in E$ for infinitely many n. Then the limit function f is also continuous at z_0.
7. Suppose $\{f_n\}$ converges uniformly to f on a compact set E, and each f_n is uniformly continuous on E. Prove that f is uniformly continuous on E. May compactness be omitted from the hypothesis?
8. If $\sum_{n=1}^{\infty} f_n(z)$ converges uniformly on E, show that $\{f_n(z)\}$ converges uniformly to zero on E. Is the converse true?
9. Let $0 < r < 1$ and $E = \{z : |z| \le r\} \cup \{z : r \le z \le 1, \ z \in \mathbb{R}\}$. Show that $\sum_{n=1}^{\infty} (-1)^n z^n/n$ converges uniformly, but not absolutely, on E.
10. Find where the following sequences converge pointwise and where uniformly.

    (a) $\dfrac{z}{z^2 + n^2}$ (b) ze^{-nz} (c) $\dfrac{e^{nz}}{n}$ (d) $\dfrac{1}{1 + z^n}$.

11. In what regions are the following series uniformly convergent? Absolutely convergent?

    (a) $\displaystyle\sum_{n=1}^{\infty} (1 - z)z^n$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{z^2}{(1 + z^2)^n}$

    (c) $\displaystyle\sum_{n=1}^{\infty} \frac{2z}{n^2 - z^2}$ (d) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{1 + z^n}$

12. Show that the series $\sum_{n=0}^{\infty} 1/[(z + n)(z + n + 1)]$ converges to $1/z$ for $z \in \mathbb{C}\backslash\{0, -1, -2, \ldots\}$.

6.3 Maclaurin and Taylor Series

Unfortunately, knowing that an arbitrary series $\sum_{n=1}^{\infty} f_n(z)$ converges (or diverges) at some point $z = z_0$ gives no information about the series at other points. However, we can specialize the sequence $\{f_n(z)\}$ to obtain a class of functions for which the behavior at a point always determines properties in a region. This class will play an important role in the theory of analytic functions.

Let $f_n(z) = a_n(z-b)^n$, where a_n and b are complex constants. The resultant expression,

$$a_0 + \sum_{n=1}^{\infty} a_n(z-b)^n = \sum_{n=0}^{\infty} a_n(z-b)^n, \tag{6.11}$$

is called a *power series* in $z-b$. When $b=0$, (6.11) reduces to

$$\sum_{n=0}^{\infty} a_n z^n, \tag{6.12}$$

a power series in z. Our efforts will be focused on properties of power series defined by (6.12). Upon substituting $z-b$ for z, it becomes a simple matter to translate these properties to those of series defined by (6.12). We state and prove a result which provides a "qualitative" behavior of a convergent power series.

Theorem 6.38. *Suppose the power series $\sum_{n=0}^{\infty} a_n z^n$ converges at a point $z=z_0$. Then $\sum_{n=0}^{\infty} |a_n|\,|z|^n$ converges for $|z|<|z_0|$ (that is, $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely for $|z|<|z_0|$).*

Proof. Suppose that $\sum_{n=0}^{\infty} a_n z_0^n$ converges. Then $a_n z_0^n \to 0$ as $n \to \infty$. Hence there exists a constant M such that $|a_n z_0^n| \le M$ for all n, and

$$|a_n|\,|z|^n = \left| a_n z_0^n \left(\frac{z}{z_0}\right)^n \right| \le M \left|\frac{z}{z_0}\right|^n. \tag{6.13}$$

For $|z|<|z_0|$, the geometric series $\sum_{n=0}^{\infty} |z/z_0|^n$ converges. Thus, by (6.13),

$$\sum_{n=0}^{\infty} |a_n|\,|z|^n \le \sum_{n=0}^{\infty} M \left|\frac{z}{z_0}\right|^n = \frac{M}{1-|z/z_0|} \quad (|z|<|z_0|). \qquad \blacksquare$$

Corollary 6.39. *If $\sum_{n=0}^{\infty} a_n z^n$ diverges at a point $z=z_0$, then $\sum_{n=0}^{\infty} a_n z^n$ diverges for $|z|>|z_0|$.*

Proof. Suppose, on the contrary, that the series $\sum_{n=0}^{\infty} a_n z^n$ converges for some point z_1 with $|z_1|>|z_0|$. Then, by Theorem 6.38, $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely for $|z|<|z_1|$. In particular, this would imply the absolute convergence of $\sum_{n=0}^{\infty} a_n z_0^n$, contradicting our assumption. $\blacksquare$

Corollary 6.40. *If $\sum_{n=0}^{\infty} a_n z^n$ converges for all real values of z, then the series also converges for all complex values.*

Proof. Suppose $\sum_{n=0}^{\infty} a_n z_0^n$ diverges for some complex value z_0. By Corollary 6.39, $\sum_{n=0}^{\infty} a_n R^n$ diverges for $R>|z_0|$, contradicting our assumption. $\blacksquare$

Theorem 6.38 can be used to determine precise bounds for the region in which a power series converges.

Theorem 6.41. *To every power series $\sum_{n=0}^{\infty} a_n z^n$, there corresponds an R, $0 \le R \le \infty$, for which the series*

(i) *converges absolutely in $|z| < R$ if $0 < R \le \infty$*
(ii) *converges uniformly in $|z| \le r < R$ if $0 < R \le \infty$*
(iii) *diverges for $|z| > R$ if $0 \le R < \infty$.*

Proof. For $z = 0$, the series becomes a_0 and hence, the power series converges at the origin. If the series diverges for all nonzero values of z, then clearly $R = 0$. If the series converges for some nonzero value, we let

$$S = \left\{ r : \sum_{n=0}^{\infty} |a_n|\,|z|^n \ \text{ converges for } \ |z| < r \right\},$$

and define

$$R = \begin{cases} \operatorname{lub} S \text{ if } S \text{ is bounded}, \\ \quad \infty \text{ if } S \text{ is unbounded}. \end{cases}$$

We wish to show that R, so chosen, satisfies conditions (i), (ii), and (iii).

For any point z_0, $|z_0| < R$, we can find a real number ρ such that

$$|z_0| < \rho < R. \tag{6.14}$$

The number ρ must be in S; for otherwise, R could not be its least upper bound. Hence, by (6.14) and the definition of set S, the series $\sum_{n=0}^{\infty} |a_n|\,|z_0|^n$ converges. Since z_0 was arbitrary, (i) is proved.

Next, if $R > 0$, choose r so that $0 < r < R$. Then there exists z_0 such that $r < |z_0| < R$ and the series $\sum_{n=1}^{\infty} |a_n z_0^n|$ is convergent (if $R = \infty$, this works for any r). In particular, the nth term $|a_n z_0^n|$ is bounded, say by a number K. Now for $|z| \le r$,

$$|a_n z^n| \le |a_n| r^n = |a_n z_0^n| \left(\frac{r}{|z_0|} \right)^n \le K \left(\frac{r}{|z_0|} \right)^n = K M_n.$$

Since $\sum M_n$ is a convergent geometric series, the Weierstrass M-test applies, and the series $\sum a_n z^n$ converges uniformly for $|z| \le r$. This proves (ii).

Finally, the convergence of the series at some point z_1, $|z_1| > R$, would, according to Theorem 6.38, imply absolute convergence for $|z| < |z_1|$. But then $|z_1|$ would be an element of S; this would contradict the assumption that R is an upper bound for the set S. This proves (iii). ■

The number R, defined by Theorem 6.41, is called the *radius of convergence* and the circle $|z| = R$ is often referred to as the *circle of convergence*. If R is the radius of convergence of a series, then the disk $|z| < R$ is called the *disk/domain of convergence* of the corresponding power series. The radius of convergence depends only on the tail of the series. If we alter a finite number of coefficients a_n of the series, the radius of convergence remains unchanged. Further, a power series always converges inside and diverges outside the circle $|z| = R$. But there is no general principle regarding its behavior *on* the circle. Therefore, if it is required to find the convergence of a power series on its circle of convergence, then this has to be investigated separately because it may converge at all, none, or some of the points. We illustrate the last fact by the following examples:

Examples 6.42. (i) The geometric series $\sum_{n=0}^{\infty} z^n$ converges for $|z| < 1$ and diverges everywhere on $|z| = 1$, its circle of convergence. Note that the series is not uniformly convergent on the open disk $|z| < 1$.

(ii) The series $\sum_{n=1}^{\infty} z^n/n$ converges at $z = -1$ and diverges at $z = 1$. Therefore, its radius of convergence must be $R = 1$ (Why?).

(iii) The series $\sum_{n=1}^{\infty}(z^n/n^2)$ converges absolutely (and uniformly) for $|z| \leq 1$. This follows from the Weierstrass M-test with majorants $M_n = 1/n^2$. If $z = 1 + \epsilon > 1$, then $(1 + \epsilon)^n/n^2 \to \infty$ as $n \to \infty$. Hence, the series does not converge for $|z| > 1$, and the radius of convergence of the series $\sum_{n=1}^{\infty}(z^n/n^2)$ is $R = 1$.

(iv) The series $\sum_{n=1}^{\infty} n^n z^n$ cannot converge for any nonzero complex values because $|nz|^n = |z|^n n^n \to \infty$ $(z \neq 0)$. Therefore, $R = 0$.

(v) The series $\sum_{n=1}^{\infty}(z^n/n^n)$ converges everywhere. To see this, choose $z = z_0$. Then for $N > |z_0|$,

$$\sum_{n=N}^{\infty} \left|\frac{z_0^n}{n^n}\right| < \sum_{n=N}^{\infty} \left|\frac{z_0}{N}\right|^n = \frac{|z_0/N|^N}{1 - |z_0/N|},$$

and the absolute convergence of $\sum_{n=1}^{\infty}(z_0^n/n^n)$ follows. Since z_0 was arbitrary, $R = \infty$. By Theorem 6.41, the series converges uniformly on all compact subsets of the plane. ●

Example 6.43. Let us discuss the convergence of the series

$$\sum_{n=-\infty}^{\infty} 3^{-|n|} z^{2n}.$$

According to the discussion on geometric series, we may rewrite the given series as

$$\begin{aligned}\sum_{n=-\infty}^{\infty} 3^{-|n|} z^{2n} &= \sum_{n=0}^{\infty} 3^{-n} z^{2n} + \sum_{n=-\infty}^{-1} 3^n z^{2n} \\ &= \sum_{n=0}^{\infty} \left(\frac{z^2}{3}\right)^n + \sum_{n=1}^{\infty} (3z^2)^{-n}.\end{aligned}$$

Note that the first series on the right converges when $|z|^2 < 3$, i.e., when $|z| < \sqrt{3}$ and that it diverges for all z with $|z| \geq \sqrt{3}$. Similarly, the second series on the right converges when $|3z^2| > 1$, i.e., when $|z| > 1/\sqrt{3}$ and that it diverges for $|z| \leq 1/\sqrt{3}$. It follows that

$$\sum_{n=-\infty}^{\infty} 3^{-|n|} z^{2n} = \frac{1}{1 - z^2/3} + \frac{1/(3z^2)}{1 - (1/3z^2)} = \frac{8z^2}{(3 - z^2)(3z^2 - 1)}$$

for $1/\sqrt{3} < |z| < \sqrt{3}$, whereas the series diverges for all remaining z. ●

More generally, Theorem 6.41 shows that a power series $\sum_{n=0}^{\infty} a_n (z-b)^n$ either converges absolutely in $\mathbb{C}$ or only at the origin or else there exists an $R > 0$, for which the series converges absolutely in $|z - b| < R$ and diverges for $|z - b| > R$. If $0 \leq r < R$ (if $r > 0$ for $R = \infty$), then the series converges uniformly in $|z - b| \leq r$.

Thus far the radii of convergence for different power series have been determined only by the sometimes cumbersome method of testing distinct points for convergence and divergence and applying Theorem 6.38. But a power series is defined by its coefficients, and it is these coefficients alone that determine its radius of convergence.

Theorem 6.44. (Cauchy–Hadamard) *The power series $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R, where $1/R = \limsup_{n\to\infty} |a_n|^{1/n}$ (Here we observe the conventions $1/0 = \infty$ and $1/\infty = 0$).*

Proof. For any point $z_0 \neq 0$, we have

$$\limsup_{n\to\infty} |a_n z_0^n|^{1/n} = \limsup_{n\to\infty} |a_n|^{1/n} |z_0| = \frac{1}{R}|z_0|.$$

According to Theorem 6.13, the series $\sum_{n=0}^{\infty} a_n z_0^n$ converges absolutely when $|z_0| < R$ and diverges when $|z_0| > R$. In view of Theorem 6.41, the radius of convergence is R.

When $R = \infty$, the series converges everywhere; and when $R = 0$, the series converges only at $z = 0$. ■

Most often the following result suffices to examine the convergence properties of the power series.

Corollary 6.45. *The radius of convergence R of the power series $\sum_{n=0}^{\infty} a_n z^n$ is determined by*

(a) $\displaystyle \frac{1}{R} = \lim_{n\to\infty} |a_n|^{1/n}$ (b) $\displaystyle \frac{1}{R} = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$

provided these limits exist.

Examples 6.46. (i) Consider $\sum_{n=1}^{\infty} a_n z^n$, $4a_n = n/(4n^2+1)$. Then the ratio test given by Corollary 6.16 is applicable, as $a_n \neq 0$ for each $n \in \mathbb{N}$. We have

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n+1}{4(n+1)^2+1}\,\frac{4n^2+1}{n} \to 1 \text{ as } n \to \infty.$$

So, the series converges for $|z| < 1$ and diverges for $|z| > 1$.

(ii) Consider $\sum_{n=0}^{\infty} \frac{1}{(3+i)^n} z^{3n}$. Note that the ratio test is not directly applicable. However, we may think of this as a series in a new variable z^3 rather than in z:

$$\sum_{n=0}^{\infty} \frac{1}{(3+i)^n} w^n, \quad w = z^3.$$

Now, we can apply the ratio test to this new series. It follows that

$$\frac{1}{R} = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{(3+i)^n}{(3+i)^{n+1}}\right| = \frac{1}{\sqrt{10}}, \quad \text{i.e., } R = \sqrt{10}$$

and so the new series converges for $|w| < \sqrt{10}$ and diverges for $|w| > \sqrt{10}$. This is equivalent to saying that the original series converges for $|z| < 10^{1/6}$ and diverges for $|z| > 10^{1/6}$.

(iii) Consider $\sum_{n=1}^{\infty} \frac{n+1}{n!} z^{n^3}$. Again the ratio test is not applicable. But we may fix $z \neq 0$ and let

$$a_n = \frac{n+1}{n!} z^{n^3}.$$

Then,

$$\frac{a_{n+1}}{a_n} = \frac{(n+2)z^{(n+1)^3}}{(n+1)!} \times \frac{n!}{(n+1)z^{n^3}} = \frac{(n+2)}{(n+1)^2} z^{(n+1)^3-n^3}$$

and so

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\frac{(n+2)}{(n+1)^2}|z|^{3n(n+1)+1}$$

which is less than 1 provided $|z| < 1$, showing that the given series converges for $|z| < 1$ and diverges $|z| > 1$.

(iv) Finally, we consider $\sum_{n=0}^{\infty} a_n z^n$, where

$$a_n = \begin{cases} i2^n & \text{for } n \text{ even} \\ -3^n & \text{for } n \text{ odd}. \end{cases}$$

Then

$$|a_n|^{1/n} = \begin{cases} 2 & \text{for } n \text{ even} \\ 3 & \text{for } n \text{ odd}, \end{cases}$$

showing that $|a_n|^{1/n}$ oscillates finitely. So, by Theorem 6.44 we see that

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n} = 3, \quad \text{i.e.,} \quad R = 1/3,$$

and the series converges for $|z| < 1/3$ and diverges for $|z| > 1/3$. ●

Example 6.47. Suppose that the power series $\sum_{n=0}^{\infty} a_n z^n$ has a positive radius of convergence. Then the series $\sum_{n=1}^{\infty} \frac{a_n}{n^n} z^n$ is entire.

To see this, let $R > 0$ be the radius of convergence of $\sum_{n=1}^{\infty} a_n z^n$. Then $\sum_{n=0}^{\infty} a_n r^n$ converges, where $0 < r < R$. In particular, $\{a_n r^n\}$ is a bounded sequence. Thus, there exist an M such that $|a_n r^n| \leq M$ for all $n \geq 0$. Now

$$\left|\frac{a_n}{n^n}\right|^{1/n} \leq \frac{M^{1/n}/r}{n} \to 0 \quad \text{as } n \to \infty$$

which, according to the root test, shows that $\sum_{n=1}^{\infty} a_n n^{-n} z^n$ is entire. ■

Example 6.48. It is easier to show that the radius of convergence of $\sum_{n=0}^{\infty}(5+(-1)^n)z^n$ is 1. Indeed, we note that

$$a_n = 5 + (-1)^n = \begin{cases} 4 & \text{if } n \text{ is odd} \\ 6 & \text{if } n \text{ is even.} \end{cases}$$

Clearly, the ratio test is not applicable. On the other hand, as $|z|^{1/n} \to 1$ for $z \neq 0$, it follows that

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n} = 1$$

which gives $R = 1$. How can we find the sum $f(z)$ of the given series? Rewrite the given series as

$$f(z) = 5\sum_{n=0}^{\infty} z^n + \sum_{n=0}^{\infty} (-1)^n z^n.$$

Note that both the series on the right are known to converge for $|z| < 1$, and diverge for $|z| \geq 1$. The sum is then given by

$$f(z) = \frac{5}{1-z} + \frac{1}{1+z} = \frac{2(3+2z)}{1-z^2}. \qquad \bullet$$

Example 6.49. Let us find the radius of convergence of the series

$$\sum_{n=0}^{\infty} \sin(n\pi/4) z^n$$

and also its sum $f(z)$. To do this, we set $a_n = \sin(n\pi/4)$. As $|a_n| \leq 1$ for all $n \geq 0$, we have

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n} \leq 1, \quad \text{i.e., } R \geq 1.$$

On the other hand $a_n = \pm 1$ for infinitely many n which shows that $1/R \geq 1$, i.e. $R \leq 1$. Hence, $R = 1$. Note that

$$a_{2(4k+1)} = \sin(2k\pi + \pi/2) = 1 \quad \text{and} \quad a_{2(4k+3)} = \sin(2k\pi + 3\pi/2) = -1.$$

To find the sum, we easily compute that

$$a_n = \begin{cases} 0 & \text{if } n = 4k \\ (-1)^k/\sqrt{2} & \text{if } n = 4k+1 \\ (-1)^k & \text{if } n = 4k+2 \\ (-1)^k/\sqrt{2} & \text{if } n = 4k+3 \end{cases}, \quad k \in \mathbb{N}_0,$$

and therefore, we may rewrite the given series in the form

$$\begin{aligned} f(z) &= \frac{z}{\sqrt{2}} + z^2 + \frac{z^3}{\sqrt{2}} + 0 - \frac{z^5}{\sqrt{2}} - z^6 - \frac{z^7}{\sqrt{2}} - 0 + \cdots \\ &= \left(\frac{z}{\sqrt{2}} + z^2 + \frac{z^3}{\sqrt{2}}\right)(1 - z^4 + z^8 - \cdots) \\ &= \frac{z(1+\sqrt{2}z+z^2)}{\sqrt{2}(1+z^4)}. \end{aligned}$$

●

Suppose the power series $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R. We wish to characterize as fully as possible the behavior of the function $f(z)$, defined by the power series, at points interior to its circle of convergence. Implicit in our work is the continuity of $f(z)$ for $|z| < R$. To show continuity at an arbitrary point z_0, $|z_0| < R$, we note (by Theorem 6.41) that the sequence $S_n(z) = \sum_{k=0}^{n} a_k z^k$ converges uniformly to $f(z)$ in the disk $|z| \le r = |z_0|$. Since $S_n(z)$ is continuous at $z = z_0$ for every n, Theorem 6.26 may be applied to establish continuity of the limit function $f(z)$ at $z = z_0$.

The differentiability of $f(z)$ is not so straightforward. We might expect the derivatives of a sequence of uniformly convergent differentiable functions to converge to a differentiable function. However, consider the sequence $\{f_n(z)\}$, where

$$f_n(z) = (\sin nz)/\sqrt{n}.$$

Although $\{f_n(z)\}$ converges uniformly on the real line, the sequence $\{f_n'(z)\}$, where $f_n'(z) = \sqrt{n}\cos nz$, converges for no real values.

Fortunately, no such pathological behavior can occur for the sequence of partial sums of a power series. In a sense, a power series may be thought of as a polynomial of infinite degree; indeed, a polynomial can be defined as a power series in which all but a finite number of coefficients are zero. The derivative of a polynomial $P_n(z) = \sum_{k=0}^{n} a_k z^k$ is

$$P_n'(z) = \sum_{k=1}^{n} k a_k z^{k-1}.$$

We will prove a similar result for power series. But first we need the following:

Lemma 6.50. *The two power series*

$$\sum_{n=0}^{\infty} a_n z^n \quad \text{and} \quad \sum_{n=1}^{\infty} n a_n z^n$$

have the same radius of convergence.

Proof. Using properties of the limit superior (Exercise 6.19(7)), we have

$$\limsup_{n\to\infty} |na_n|^{1/n} = \lim_{n\to\infty} n^{1/n} \limsup_{n\to\infty} |a_n|^{1/n} = \limsup_{n\to\infty} |a_n|^{1/n}.$$

The result now follows from Theorem 6.44. ■

Theorem 6.51. *If a function $f(z)$ is the pointwise limit of a power series $\sum_{n=0}^{\infty} a_n z^n$ in $|z| < R$, then $f(z)$ is analytic for $|z| < R$, with*

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}.$$

Proof. Given z_0, $|z_0| < R$, we will show that

$$\left| \frac{f(z) - f(z_0)}{z - z_0} - \sum_{n=1}^{\infty} n a_n z_0^{n-1} \right| < \epsilon$$

whenever $|z - z_0| < \delta = \delta(\epsilon, z_0)$ (see Fig. 6.2).

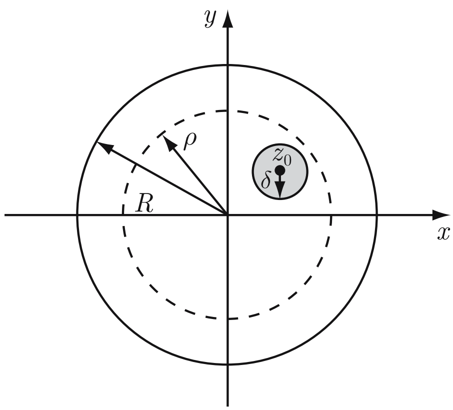


Figure 6.2.

For z sufficiently close to z_0, there is a real number ρ satisfying the inequalities

$$|z_0| \leq \rho, \quad |z| \leq \rho \quad (\rho < R). \tag{6.15}$$

For any integer N, we write

$$P_N(z) = \sum_{n=0}^{N} a_n z^n$$

so that $P'_N(z) = \sum_{n=1}^{N} n a_n z_0^{n-1}$. Therefore, we have

$$\begin{aligned}
&\left| \frac{f(z) - f(z_0)}{z - z_0} - \sum_{n=1}^{\infty} n a_n z_0^{n-1} \right| \qquad (6.16)\\
&\quad = \left| \frac{P_N(z) - P_N(z_0)}{z - z_0} - P'_N(z_0) + \sum_{n=N+1}^{\infty} a_n \left(\frac{z^n - z_0^n}{z - z_0} - n z_0^{n-1} \right) \right| \\
&\quad \le \left| \frac{P_N(z) - P_N(z_0)}{z - z_0} - P'_N(z_0) \right| + \left| \sum_{n=N+1}^{\infty} a_n \left(\frac{z^n - z_0^n}{z - z_0} - n z_0^{n-1} \right) \right| .
\end{aligned}$$

Denote the last two expressions by A_1 and A_2, respectively. We shall first choose N large enough so that $A_2 < \epsilon/2$, and then choose δ small enough so that $A_1 < \epsilon/2$. From (6.15),

$$\begin{aligned}
\left| \frac{z^n - z_0^n}{z - z_0} - n z_0^{n-1} \right| &\le |z^{n-1} + z^{n-2} z_0 + \cdots + z_0^{n-1}| + n|z_0|^{n-1} \qquad (6.17)\\
&\le |z|^{n-1} + |z|^{n-2}|z_0| + \cdots + |z_0|^{n-1} + n|z_0|^{n-1} \\
&\le n\rho^{n-1} + n\rho^{n-1} = 2n\rho^{n-1}.
\end{aligned}$$

According to the lemma, $\sum_{n=1}^{\infty} n|a_n|\rho^{n-1}$ converges and, by the Cauchy criterion, an integer N may be found such that

$$\sum_{n=N+1}^{\infty} n|a_n|\rho^{n-1} < \frac{\epsilon}{4}, \quad \text{i.e.,} \quad A_2 < \epsilon/2. \qquad (6.18)$$

The inequality

$$A_1 < \frac{\epsilon}{2} \quad \text{for} \quad |z - z_0| < \delta \qquad (6.19)$$

is a consequence of the differentiability of the polynomial $P_N(z)$ at $z = z_0$. If we combine (6.18) and (6.19), the result follows from (6.16). ■

Remark 6.52. A power series may always be written as a polynomial plus a "tail". The essence of this proof consisted of showing the tail to be inconsequential. ●

Remark 6.53. Theorem 6.51 says that every function defined by its power series is analytic inside its radius of convergence. In Chapter 8, the converse will also be proved. That is, every function analytic in a disk may be expressed as a power series. ●

An examination of Theorem 6.51 reveals that much more has been proved than was originally intended. The power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ was shown to have derivative

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}$$

which, itself, is a power series having the same radius of convergence as $f(z)$. Thus, Theorem 6.51 may be applied repeatedly to obtain

$$\begin{aligned} f^{(k)}(z) &= \sum_{n=k}^{\infty} n(n-1)\cdots(n-(k-1))a_n z^{n-k} \\ &= k!a_k + \frac{(k+1)!}{1!}a_{k+1}z + \frac{(k+2)!}{2!}a_{k+2}z^2 + \cdots, \end{aligned} \tag{6.20}$$

which is valid inside the circle of convergence of $f(z)$. Setting $z = 0$ in (6.20) we see that the coefficients a_k are related to the sum function $f(z)$ of the power series through the expression

$$f^{(k)}(0) = k!a_k, \quad \text{i.e., } a_k = \frac{f^{(k)}(0)}{k!} \tag{6.21}$$

for $k = 0, 1, 2, \ldots$. Here we have used the conventions $f^{(0)}(z) = f(z)$ and $0! = 1$. We sum up these results as

Theorem 6.54. *If a function $f(z)$ is the pointwise limit of a power series in some neighborhood of the origin, then $f(z)$ has derivatives of all orders at each point interior to the circle of convergence of $f(z)$. Furthermore, the coefficients of the power series are uniquely determined and are related to the derivatives of $f(z)$ at the origin by* (6.21).

The representation

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n$$

is called the *Maclaurin series* expansion of $f(z)$.

Example 6.55. Let us sum the series $\sum_{n=1}^{\infty} n(n+3)z^n$ for $|z| < 1$. To do this, for $z \neq 0$, the geometric series shows that

$$\frac{z^3}{1-z} = \sum_{n=0}^{\infty} z^{n+3}$$

from which one obtains that

$$\frac{1}{z^2}\left(\frac{z^3}{1-z}\right)' = \sum_{n=0}^{\infty} (n+3)z^n.$$

Differentiating the left-hand side and then multiplying the resulting expression by z would yield the desired sum. ●

All of our results on power series may easily be reworded to accommodate expansions in powers of $z - b$, where b is any complex number. For instance, the series $\sum_{n=0}^{\infty} a_n(z-b)^n$ converges absolutely to an analytic function $f(z)$ inside the circle $|z-b| = R$, where $R^{-1} = \limsup_{n\to\infty} |a_n|^{1/n}$. Moreover, the *Taylor series* expansion

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(b)}{n!}(z-b)^n$$

is valid for $|z-b| < R$.

Theorem 6.56. (see also Theorem 8.44) *Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be a power series with positive radius of convergence and $\{z_k\}_{k\geq 1}$ a sequence which converges to zero such that $z_k \neq 0$ for all $k \in \mathbb{N}$. Further assume that $f(z_k) = 0$ for $k \in \mathbb{N}$. Then $a_n = 0$ for all $n \in \mathbb{N}_0$.*

Proof. As f is analytic at $z = 0$, f is continuous at $z = 0$. We have

$$z_k \to 0 \quad \Rightarrow \quad f(z_k) \to f(0) \quad \Rightarrow \quad f(0) = 0.$$

Next consider the function $f_1(z) = \sum_{n=1}^{\infty} a_n z^{n-1}$ which has the same radius of convergence as f, and

$$0 = f(z_k) = f_1(z_k) z_k, \quad k \in \mathbb{N}.$$

Since $z_k \neq 0$, $f_1(z_k) = 0$ for $k \in \mathbb{N}$. Hence $a_1 = f_1(0) = 0$. Continuing this process we obtain the desired result. ∎

Questions 6.57.

1. Suppose a power series converges at $z = z_0$ and diverges at $z = z_1$. What is the relationship between z_0 and z_1?
2. Suppose a power series converges at all the positive integers. What kind of function does it represent?
3. Can a power series $\sum_{n=0}^{\infty} a_n(z-5i)^n$ converge at $z = 0$ and diverge at $z = 1 + 7i$?
4. Is the set S, defined in Theorem 6.41, a closed set?
5. When will the regions of absolute and uniform convergence of a power series coincide?
6. How do the convergence properties of $\sum_{n=0}^{\infty} a_n z^n$ and $\sum_{n=0}^{\infty} n a_n z^n$ compare?
7. If $\sum_{n=0}^{\infty} a_n$ converges, what can be said about the radius of convergence of $\sum_{n=0}^{\infty} a_n z^n$?
8. In what ways do power series having radius of convergence $R = 0$ or $R = \infty$ differ from other power series?
9. Suppose $\{f_n\}$ is a sequence of differentiable functions, and $\{f_n'\}$ converges uniformly on a set E to a differentiable function. Must $\{f_n\}$ converge on E?

10. What can be said about the sum and product of power series?
11. What analytic functions can be shown to have power series representations?
12. Are the theorems in this section valid for real power series?

Exercises 6.58.

1. Suppose $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R. Show that the sequence $\{a_n z_0^n\}$ is unbounded if $|z_0| > R$.
2. (a) If $\sum_{n=0}^{\infty} a_n z_1^n$ converges and $\sum_{n=0}^{\infty} a_n z_2^n$ diverges, with $|z_1| = |z_2|$, show that $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence $R = |z_1|$.
   (b) If $\sum_{n=0}^{\infty} a_n$ converges and $\sum_{n=0}^{\infty} |a_n|$ diverges, show that the series $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence $R = 1$.
3. Suppose $|a_n|$ is a decreasing sequence. Show that the radius of convergence of $\sum_{n=0}^{\infty} a_n z^n$ is at least 1.
4. Suppose $\sum_{n=0}^{\infty} a_n z^n$ converges at an unbounded sequence of points. Show that the power series converges everywhere.
5. Suppose $\{a_n\}$ is a sequence of integers. Prove that $\sum_{n=0}^{\infty} a_n z^n$ is either an entire function or has radius of convergence at most one.
6. Show that a power series converges uniformly on all compact subsets interior to its circle of convergence.
7. Suppose $\lim_{n\to\infty} |a_n/a_{n+1}| = R$. Show that $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R.
8. Show that the radius of convergence of any power series $\sum_{n=0}^{\infty} a_n z^n$ is given by $R = \liminf_{n\to\infty} |a_n|^{-1/n}$.
9. Show, for any integer k, that $\sum_{n=1}^{\infty} (n^k/n^{\ln n}) z^n$ has radius of convergence $R = 1$.
10. Find the radius of convergence for

    (a) $\displaystyle\sum_{n=1}^{\infty} \frac{n^k}{a^n} z^n$ (b) $\displaystyle\sum_{n=1}^{\infty} (n^k + a^n) z^n$ (c) $\displaystyle\sum_{n=1}^{\infty} \frac{i^n - 1}{n} z^n$

    (d) $\displaystyle\sum_{n=1}^{\infty} \left(1 + \frac{1}{n}\right)^{n^2} z^n$ (e) $\displaystyle\sum_{n=0}^{\infty} \frac{n^2 + 5n + 3i^n}{2n+1} z^n$ (f) $\displaystyle\sum_{n=1}^{\infty} \frac{(n!)^2}{(2n)!} z^n$

    (g) $\displaystyle\sum_{n=1}^{\infty} n^{1/n} (z+i)^n$ (h) $\displaystyle\sum_{n=1}^{\infty} (\ln n)^n z^{n^3}$ (i) $\displaystyle\sum_{n=1}^{\infty} n^3 (z+1)^{3^n}$

    (j) $\displaystyle\sum_{n=1}^{\infty} n! z^n$ (k) $\displaystyle\sum_{n=0}^{\infty} n! z^{n!}$ (l) $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n} z^n.$

11. Find the radius of convergence for

    (a) $\displaystyle\sum_{n=1}^{\infty} \frac{z^{2n}}{4^n n^k}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{n^2 + 4n} z^{2n}$ (c) $\displaystyle\sum_{n=0}^{\infty} \frac{z^n}{2^{n^2}}$

    (d) $\displaystyle\sum_{n=0}^{\infty} \frac{z^{n^2}}{2^n}$ (e) $\displaystyle\sum_{n=0}^{\infty} \frac{2^n + 3^n}{4^n + 5^n} z^n$ (f) $\displaystyle\sum_{n=0}^{\infty} 2^n z^{n^2}$

    (g) $\displaystyle\sum_{n=0}^{\infty} \cos(n\pi/6) z^n$ (h) $\displaystyle\sum_{n=0}^{\infty} \sin(n\pi/16) z^n.$

12. Suppose $\{a_n\}$ is a complex sequence whose partial sums $\sum_{i=1}^{n} a_i$ are bounded. If $\{b_n\}$ is a real sequence that is monotonically decreasing to 0, show that $\sum_{n=1}^{\infty} a_n b_n$ converges.
13. (a) Show that $\sum_{n=1}^{\infty}(z^n/n)$ converges everywhere on the circle $|z| = 1$ except $z = 1$.
    (b) Show, for $|z_1| = 1$, that $\sum_{n=1}^{\infty}(1/n)(z/z_1)^n$ converges everywhere on the unit circle except $z = z_1$.
    (c) Suppose $|z_1| = |z_2| = \cdots = |z_p| = 1$. Show that

$$\sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{1}{z_1^n} + \cdots + \frac{1}{z_p^n}\right) z^n$$

    converges everywhere on the unit circle except $z_1, z_2, \ldots, z_p$.
14. Write Taylor expansions for the polynomial $P(z) = z^3 + 3z^2 - 2z + 1$ in powers of

$$\text{(a)}\ \ z-1 \quad \text{(b)}\ \ z+2 \quad \text{(c)}\ \ z-i.$$

15. Show that the series

$$\sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{8^n}$$

    converges for $|z| < 2\sqrt{2}$. Find its sum.
16. Suppose that $a_n \neq 0$ for all $n \in \mathbb{N}$. If R is the radius of convergence of both the series

$$\sum_{n=1}^{\infty} a_n z^n \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{z^n}{a_n},$$

    then show that $R = 1$.
17. Sum the series $\sum_{n=0}^{\infty} \cos(n\pi/3) z^n$.

6.4 Operations on Power Series

Our study of power series has revolved around the following three questions:

(i) For what values of z does $\sum_{n=0}^{\infty} a_n(z-b)^n$ converge?
(ii) What properties may be attributed to $f(z) = \sum_{n=0}^{\infty} a_n(z-b)^n$ at points where the series converges?
(iii) Under what conditions may a function $f(z)$ be represented by a power series in some neighborhood of a point?

The first two questions are almost completely solved. The series

$$\sum_{n=0}^{\infty} a_n(z-b)^n$$

either converges everywhere, only at $z = b$, or there exists a circle for which the series converges absolutely inside and diverges outside. The function

$$f(z) = \sum_{n=0}^{\infty} a_n (z-b)^n$$

is analytic with derivatives of all orders inside the circle of convergence. Only its behavior on the circle remains a mystery.

The third question has not yet been properly dealt with. We know that $f(z)$ cannot be represented by a power series in a neighborhood of a point unless it has derivatives of all orders at each point in the neighborhood. Furthermore, if $f(z)$ does have a power series representation in some neighborhood of $z = b$, then that representation is unique, and its coefficients are related to the derivatives at $z = b$ by

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(b)}{n!}(z-b)^n.$$

But we still have no criteria that will guarantee a power series development.

Consider, for example, the function

$$f(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots .$$

By the ratio test, the radius of convergence is ∞, and so $f(z)$ is entire. Alternatively, in view of Theorem 6.44, it suffices to show that $(n!)^{1/n} \to \infty$. Beginning with the inequality

$$n! \geq n(n-1) \cdots \left(n - \frac{n}{2}\right) \geq \left(\frac{n}{2}\right)^{n/2},$$

we take nth roots of both sides to obtain

$$(n!)^{1/n} \geq \left(\frac{n}{2}\right)^{1/2} \to \infty.$$

Thus, $f(z)$ is analytic everywhere. Moreover, $f(0) = 1$ and, by Theorem 6.51, $f'(z) = f(z)$ for all z.

We would like very much, at this point, to say that $f(z) = e^z$. In fact, if e^z does have a power series representation, then

$$f(z) = e^z = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n = \sum_{n=0}^{\infty} \frac{z^n}{n!}. \tag{6.22}$$

To give us even more faith in the truth of (6.22), note that the identity

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

is valid for all real x. This is proved in elementary calculus by use of Taylor's formula with remainder [T]. That is,

$$f(x) = e^x = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + R_n(x),$$

where $R_n(x)$, defined in terms of a real integral, approaches zero as n approaches infinity.

The proof of (6.22) for complex z will be postponed until the theory of complex integration has been developed, but even the most impatient reader will find it worth the wait. Not only will we prove that all the familiar real power series identities like

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \cdots,$$

$$\sin x = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^{2n-1}}{(2n-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots,$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

remain valid in the complex plane, but also that there is a converse to Theorem 6.51; namely, that *every* analytic function admits a power-series expansion. In particular, if $f(z)$ is an entire function, then its Taylor series representation

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(b)}{n!} (z-b)^n$$

is valid for all complex b and z.

This becomes even more striking in view of the absence of a real variable analog.

Example 6.59. We remark that the function $f(x) = x|x|$ is differentiable for all real x, but cannot be expanded in a Maclaurin (real) series because $f''(0)$ does not exist. ●

Example 6.60. The function $f(x) = 1/(1+x^2)$ has derivatives of all orders for all real x, although the Maclaurin expansion

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = 1 - x^2 + x^4 - x^6 + \cdots$$

is valid only in the real interval $(-1, 1)$. There appears to be nothing in the nature of the function to account for this restriction. But replacing x by the complex variable z, we see that the function $f(z) = 1/(1+z^2)$ is not analytic at $z = \pm i$. This prevents a Maclaurin series from converging outside the circle $|z| = 1$. In particular, for real values of z the series cannot converge outside the real interval $[-1, 1]$. ●

Example 6.61. The function

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

has derivatives of all orders for all real values. Since $f^{(n)}(0) = 0$ for every integer n, we have $\sum_{n=0}^{\infty}(f^{(n)}(0)/n!)x^n \equiv 0$. Hence, the Maclaurin series represents the function only at the origin. ●

Returning, once again, to functions of a complex variable, the sum of two polynomials of degree n is a polynomial of degree at most n and is formed by adding coefficients termwise. That is,

$$\sum_{k=0}^{n} a_k z^k + \sum_{k=0}^{n} b_k z^k = \sum_{k=0}^{n} (a_k + b_k) z^k.$$

The product of two polynomials of degree n is a polynomial of degree $2n$, but the relationship between coefficients is not as simple. We have

$$\begin{aligned}(a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n)&(b_0 + b_1 z + b_2 z^2 + \cdots + b_n z^n) \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0) z + (a_0 b_2 + a_1 b_1 + a_2 b_0) z^2 \\ &\qquad + \cdots + a_n b_n z^{2n}.\end{aligned}$$

More concisely,

$$\left(\sum_{k=0}^{n} a_k z^k\right)\left(\sum_{k=0}^{n} b_k z^k\right) = \sum_{k=0}^{2n} c_k z^k, \quad c_k = \sum_{m=0}^{k} a_m\, b_{k-m}.$$

If two functions are known to have power series representations, then information about their sum and product can be obtained.

Theorem 6.62. *Suppose* $f(z) = \sum_{n=0}^{\infty} a_n z^n$ *and* $g(z) = \sum_{n=0}^{\infty} b_n z^n$ *have radii of convergence* R_1 *and* R_2*, respectively. Then* $f(z) + g(z)$ *and* $f(z)g(z)$ *have power series representations whose radius of convergence is at least* $R = \min\{R_1, R_2\}$.

Proof. Set $S_n(z) = \sum_{k=0}^{n} a_k z^k$ and $T_n(z) = \sum_{k=0}^{n} b_k z^k$. Then

$$S_n(z) + T_n(z) = \sum_{k=0}^{n} (a_k + b_k) z^k \quad \text{and} \quad S_n(z) T_n(z) = \sum_{k=0}^{2n} c_k z^k,$$

where $c_k = \sum_{m=0}^{k} a_m b_{k-m}$. For any point z_0, $|z_0| < R$, we have

$$\lim_{n\to\infty} S_n(z_0) = f(z_0) \quad \text{and} \quad \lim_{n\to\infty} T_n(z_0) = g(z_0).$$

Hence

$$\lim_{n\to\infty}(S_n(z_0)+T_n(z_0)) = f(z_0)+g(z_0) = \sum_{n=0}^{\infty}(a_n+b_n)z_0^n$$

and

$$\lim_{n\to\infty} S_n(z_0)T_n(z_0) = f(z_0)g(z_0) = \sum_{n=0}^{\infty} c_n z_0^n.$$

Since z_0 was arbitrary, both sides converge for $|z|<R$. ■

Remark 6.63. The radii of convergence for $\sum_{n=0}^{\infty}(a_n+b_n)z^n$ and $\sum_{n=0}^{\infty} c_n z^n$ may actually be greater than $\min\{R_1, R_2\}$. If $a_n \equiv 1$, $b_n \equiv -1$, then $R_1 = R_2 = 1$; but

$$\sum_{n=0}^{\infty}(a_n+b_n)z^n = \sum_{n=0}^{\infty}(1-1)z^n \equiv 0,$$

and the series converges for all values of z. If

$$a_n = \begin{cases} 2 & \text{if } n=0 \\ 2^n & \text{if } n\geq 1, \end{cases} \quad \text{and} \quad b_n = \begin{cases} -1 & \text{if } n=0 \\ 1 & \text{if } n\geq 1, \end{cases}$$

then $R_1 = \frac{1}{2}$ and $R_2 = 1$. However, $c_0 = a_0 b_0 = -2$ and for $n \geq 1$,

$$\begin{aligned} c_n = \sum_{k=0}^{n} a_k b_{n-k} &= a_0 b_n + a_n b_0 + \sum_{k=1}^{n-1} a_k b_{n-k} \\ &= 2 - 2^n + \sum_{k=1}^{n-1} 2^k \\ &= 2 - 2^n + \frac{2-2^n}{1-2} = 0. \end{aligned}$$

Therefore, $\sum_{n=0}^{\infty} c_n z^n = c_0 = -2$, and the series converges for all z. Note that

$$f(z) = 2 + \sum_{n=1}^{\infty} 2^n z^n = 2 + \frac{2z}{1-2z} = \frac{2(1-z)}{1-2z} \quad \left(|z| < \frac{1}{2}\right)$$

and

$$g(z) = -1 + \sum_{n=1}^{\infty} z^n = -1 + \frac{z}{1-z} = -\frac{1-2z}{1-z} \quad (|z|<1). \quad \bullet$$

The only function, essentially, whose Maclaurin expansion we know as yet in "closed form" is the geometric series

$$f(z) = \frac{1}{1-z} = \sum_{n=0}^{\infty} z^n \quad (|z|<1).$$

For any nonzero complex number a, this leads to the identity

$$\frac{1}{a-z} = \frac{1}{a(1-z/a)} = \frac{1}{a}\sum_{n=0}^{\infty}\left(\frac{z}{a}\right)^n = \sum_{n=0}^{\infty}\frac{1}{a^{n+1}}z^n \quad (|z| < |a|).$$

Also, for any two distinct complex numbers a and b, we have

$$\begin{aligned}\frac{1}{(z-a)(z-b)} &= \frac{1}{a-b}\left(\frac{1}{z-a} - \frac{1}{z-b}\right)\\ &= \frac{1}{a-b}\left[\sum_{n=0}^{\infty}\left(\frac{1}{b^{n+1}} - \frac{1}{a^{n+1}}\right)z^n\right]\end{aligned}$$

valid in the region

$$|z| < R = \min\{|a|, |b|\}. \tag{6.23}$$

A Maclaurin expansion for $1/(1-z)^2$ can be found by two different methods.

Method 1: Setting $a_n \equiv b_n \equiv 1$ in the proof of Theorem 6.54, we have

$$\begin{aligned}\frac{1}{(1-z)^2} &= \frac{1}{1-z}\cdot\frac{1}{1-z}\\ &= \left(\sum_{n=0}^{\infty}z^n\right)\left(\sum_{n=0}^{\infty}z^n\right)\\ &= \sum_{n=0}^{\infty}(n+1)z^n \quad (|z| < 1).\end{aligned}$$

Method 2: By Theorem 6.51 and the geometric series defined above,

$$f'(z) = \frac{1}{(1-z)^2} = \sum_{n=1}^{\infty}nz^{n-1} \quad (|z| < 1). \tag{6.24}$$

It is usually much easier to decide whether or not a given series converges than it is to find the value of a known convergent series. For instance, any student of elementary calculus can show that the series $\sum_{n=0}^{\infty}(1/n^3)$ converges, but the finest mathematicians in the world have not yet developed methods to find its sum. However, all is not lost; the closed form of the geometric series does enable us to find the value of many series.

Let us find the sum of $\sum_{n=1}^{\infty} n/2^n$. To do this, by (6.24), we have

$$zf'(z) = \frac{z}{(1-z)^2} = \sum_{n=1}^{\infty}nz^n \quad (|z| < 1). \tag{6.25}$$

Letting $z = \frac{1}{2}$ in (6.25), we obtain

$$\sum_{n=1}^{\infty}\frac{n}{2^n} = \frac{\frac{1}{2}}{(1-\frac{1}{2})^2} = 2.$$

It is interesting to note that $\sum_{n=0}^{\infty} 2^{-n} = \sum_{n=0}^{\infty} n2^{-n}$. The apparent paradox may be resolved by noting that the first term on the right vanishes.

Example 6.64. Let us find the sum of $\sum_{n=1}^{\infty}(n^2/3^n)$. Differentiating (6.25) and multiplying by z, we have

$$\frac{z+z^2}{(1-z)^3} = \sum_{n=1}^{\infty} n^2 z^n \quad (|z| < 1). \tag{6.26}$$

Setting $z = \frac{1}{3}$ in (6.26) leads to

$$\sum_{n=1}^{\infty} \frac{n^2}{3^n} = \frac{(\frac{1}{3}) + (\frac{1}{3})^2}{(1-\frac{1}{3})^3} = \frac{3}{2}.$$

●

The method employed in these examples may be used to evaluate any series of the form

$$\sum_{n=1}^{\infty} \frac{n^k}{z_0^n} \quad (k \;\text{ a positive integer,}\;\; |z_0| > 1).$$

It is sometimes possible to obtain, in closed form, a power series whose coefficients are defined recursively.

Example 6.65. The Fibonacci sequence is defined by

$$a_{n+2} = a_{n+1} + a_n \;\text{ for all } n \ge 0,$$

with $a_0 = 0$ and $a_1 = 1$. Suppose $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$. Then

$$\begin{aligned} f(z) &= z + \sum_{n=0}^{\infty} a_{n+2} z^{n+2} \\ &= z + \sum_{n=0}^{\infty} (a_{n+1} + a_n) z^{n+2} \\ &= z + z\sum_{n=0}^{\infty} a_{n+1} z^{n+1} + z^2 \sum_{n=0}^{\infty} a_n z^n \\ &= z + zf(z) + z^2 f(z). \end{aligned}$$

Solving, we obtain

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

The above manipulations are valid only at points where the series converges. The roots of the denominator of $f(z)$ are $z = (1 \pm \sqrt{5})/2$. By (6.23), the radius of convergence of $f(z)$ is seen to be $(\sqrt{5}-1)/2$.

The geometric series may also be manipulated to obtain Taylor series expansions. For example, we know that for any complex number b, $b \neq 1$, the identity

$$\frac{1}{1-z} = \frac{1}{(1-b)[1-(z-b)/(1-b)]} = \sum_{n=0}^{\infty} \frac{1}{(1-b)^{n+1}}(z-b)^n$$

is valid whenever $|z-b| < |1-b|$. Hence, $f(z) = 1/(1-z)$ has a Taylor expansion about every point except $z = 1$. The reader may wish to check that

$$\frac{f^{(n)}(b)}{n!} = \frac{1}{(1-b)^{n+1}} \quad \text{for every} \quad n. \qquad \bullet$$

Questions 6.66.

1. What are the differences between real and complex power series?
2. Does the quotient of two power series have a power series representation?
3. Suppose $f(z)$ and $g(z)$ have power series representations. What can be said about $f(g(z))$?
4. Suppose $f(z)$ has a Maclaurin expansion with radius of convergence R. Can $f(z)$ be analytic at a point z_0, $|z_0| > R$? Can $f(z)$ be analytic everywhere on the circle $|z| = R$?
5. Suppose a function is known to have a power series representation. What operations may then be used to evaluate specific infinite series?
6. If $a_n \neq 0$, how do the radii of convergence of the series $\sum_{n=0}^{\infty} a_n z^n$ and $\sum_{n=0}^{\infty} (1/a_n) z^n$ compare?
7. If a function has a Taylor expansion about two distinct points, how will the radii of convergence of the two power series compare?
8. What can be said about power series representations for rational functions?
9. Can a power series converge in an open disk $|z| < R$ without being absolutely convergent there? What about a closed disk $|z| \leq R$?

Exercises 6.67.

1. Suppose $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R_1, $0 < R_1 < \infty$, and $\sum_{n=0}^{\infty} b_n z^n$ has radius of convergence R_2, $0 < R_2 < \infty$. Show that
   (a) $\sum_{n=0}^{\infty} a_n b_n z^n$ has radius of convergence at least $R_1 R_2$.
   (b) $\sum_{n=0}^{\infty} (a_n/b_n) z^n$ $(b_n \neq 0)$ has radius of convergence at most R_1/R_2.
   Give examples to show that inequality may hold in (a) and (b).
2. Suppose $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R, $0 < R < \infty$. Find the radii of convergence for

   $$\text{(a)} \ \sum_{n=0}^{\infty} a_n n^k z^n \quad \text{(b)} \ \sum_{n=0}^{\infty} \frac{a_n}{n^k} z^n \quad \text{(c)} \ \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n \quad \text{(d)} \ \sum_{n=0}^{\infty} a_n n! z^n.$$

   Which of these answers are different if $R = 0$ or $R = \infty$?
3. Derive the power series of $(1-z)^{-4}$ (about $a \neq 1$) from the geometric series (about a).

4. Suppose $\sum_{n=0}^{\infty} a_n z^n$ has radius of convergence R. Give examples in which $\sum_{n=0}^{\infty} a_n z^{kn}$ (k a positive integer) and $\sum_{n=0}^{\infty} a_n z^{n^2}$ have radius of convergence R, and radius of convergence greater than R.
5. (a) Suppose $f(z)$ can be expanded in a Taylor series about the point $z = a$. Show that $f(z)$ is entire if $|f^{(n)}(a)| \leq M$ for some constant M and for every integer n.
   (b) Show that $f(z)$ is entire if $|f^{(n)}(a)| \leq n^k$ for some integer k and all n.
6. Find the sum of the following series.

$$\text{(a)} \sum_{n=1}^{\infty} \frac{n^2 + 2n - 1}{3^n} \qquad \text{(b)} \sum_{n=2}^{\infty} \frac{n(3^n - 2^n)}{6^n} \qquad \text{(c)} \sum_{n=1}^{\infty} \frac{5in}{(1+i)^n}.$$

7. Find the radius of convergence of the series $\sum_{n=1}^{\infty} \left(\frac{1}{n^2} + (-3)^n\right) z^n$.
8. Suppose that $a_n + A a_{n-1} + B a_{n-2} = 0$ $(n = 2, 3, 4, \dots)$. Show that

$$\sum_{n=0}^{\infty} a_n z^n = \frac{a_0 + (a_1 + a_0 A) z}{1 + Az + Bz^2}$$

at all points where the power series converges. What is the radius of convergence?
9. For the Fibonacci sequence defined by $a_{n+2} = a_{n+1} + a_n$, with $a_0 = 0$ and $a_1 = 1$, show that $a_n \leq \left(2/(\sqrt{5} - 1)\right)^n$ for every n.
10. Suppose a, b, and c are distinct nonzero complex numbers. Find a Taylor series expansion for $f(z) = 1/(z-a)(z-b)$ about the point $z = c$, and determine its radius of convergence.
11. Assume that f is analytic for $|z| < r$ for some $r > 0$ and f satisfies the functional equation $f(2z) = (f(z))^2$ for all z sufficiently close to zero (which is to make sure that both z and $2z$ lie in some disk about 0 that is contained in Δ_r). Show that f can be extended to an entire function. Determine all such entire functions explicitly.

7

Complex Integration and Cauchy's Theorem

One of the most important theorems in calculus is properly named the fundamental theorem of integral calculus. On the one hand it relates integration to differentiation, and on the other hand it gives a method for evaluating integrals. In this chapter, we mainly look for a complex analog to develop a machinery of integration along arcs and contours in the complex plane. The problem, of course, is that between any two points there are an infinite variety of paths along which to integrate. The antiderivative of a complex-valued function $f(z)$ of a complex variable z is completely analogous to that for a real function; it is indeed a complex function F whose derivative is f. Cauchy's theorem, the fundamental theorem of complex integration says that for analytic functions, one path over special domains is as good as another.

7.1 Curves

We begin by recalling some properties of the Riemann integral. Suppose $f(t)$ and $g(t)$ are real-valued functions continuous on the interval $a \leq t \leq b$. Then the Riemann integrals $\int_a^b f(t)\,dt$ and $\int_a^b g(t)\,dt$ exist. Further, for any real constants c_1 and c_2, we have the linearity property

$$\int_a^b (c_1 f(t) + c_2 g(t))\,dt = c_1 \int_a^b f(t)\,dt + c_2 \int_a^b g(t)\,dt \tag{7.1}$$

and the integral inequality

$$\left| \int_a^b f(t)\,dt \right| \leq \int_a^b |f(t)|\,dt. \tag{7.2}$$

The integration of a complex-valued function of complex variable along a contour leads to results of great importance both in pure and applied sciences. Consider now the complex-valued function

$$F(t) = F_1(t) + iF_2(t),$$

where $F_1(t)$ and $F_2(t)$ are real-valued functions continuous on the interval $[a, b]$. For example, e^{it} and $(1 + 2\cos t) - 2i \sin t$ are complex-valued functions defined on every interval in $\mathbb{R}$. Obviously, as F_1 and F_2 are integrable over the interval $[a, b]$, the definite integral of $F(t)$ is defined by

$$\int_a^b F(t)\,dt = \int_a^b F_1(t)\,dt + i \int_a^b F_2(t)\,dt. \tag{7.3}$$

First, we observe that

$$\mathrm{Re} \int_a^b F(t)\,dt = \int_a^b \mathrm{Re}\, F(t)\,dt = \int_a^b F_1(t)\,dt$$

and

$$\mathrm{Im} \int_a^b F(t)\,dt = \int_a^b \mathrm{Im}\, F(t)\,dt = \int_a^b F_2(t)\,dt.$$

Many familiar rules of integration for real-valued functions can be carried over to the complex case. For instance, the linearity property expressed in (7.1) is true for complex-valued functions and complex constants. The proof consists of separating into real and imaginary parts. To prove (7.2) for continuous complex-valued functions $F(t)$ on $[a, b]$, suppose

$$\int_a^b F(t)\,dt = Re^{i\alpha} \quad (R > 0, \ -\pi < \alpha \le \pi).$$

Then

$$\int_a^b e^{-i\alpha} F(t)\,dt = e^{-i\alpha} \int_a^b F(t)\,dt = R = \left| \int_a^b F(t)\,dt \right|. \tag{7.4}$$

In view of (7.4) and properties of the real integral,

$$\begin{aligned} R = \mathrm{Re} \int_a^b e^{-i\alpha} F(t)\,dt &= \int_a^b \mathrm{Re}\, \left(e^{-i\alpha} F(t) \right)\,dt \\ &\le \int_a^b |e^{-i\alpha} F(t)|\,dt = \int_a^b |F(t)|\,dt. \end{aligned}$$

The inequality is obvious when $\int_a^b F(t)\,dt = 0$. Thus, the magnitude of an integral does not exceed the integral of the absolute value of the integrand. Later in Theorem 7.19, we show that a similar inequality holds for integration along contours.

Suppose that $f(z)$ is a complex-valued function of a complex variable z defined on a subset $\Omega \subseteq \mathbb{C}$. Suppose that z_1 and z_2 are two points in Ω. At first, we are concerned with the following:

Problem 7.1. How do we define the integral of a complex-valued function f of complex variable z from z_1 to z_2?

Our discussion thus far provides no help in dealing with Problem 7.1. Overcoming this problem will necessitate integrating along more general curves than real intervals. First, we must define the notion of a curve in $\mathbb{C}$.

A *continuous curve* (an *arc*) C in the complex plane is defined parametrically by

$$C : z(t) = x(t) + iy(t) \quad (t \in [a,b],\ a < b), \tag{7.5}$$

where $x(t)$ and $y(t)$ are real-valued, continuous functions of the real variable t. We will henceforth assume that all curves are continuous curves, so that the terms "curve" and "arc" may be used interchangeably. So, by a curve C in $\mathbb{C}$, we mean a continuous function from $[a,b]$ into $\mathbb{C}$.

A curve may have more than one parameterization. For instance,

$$z_1(t) = t \ (t \in [0,1]) \ \text{ or } \ z_2(t) = t^2 \ (t \in [0,1])$$

represents the interval $[0,1]$. A natural ambiguity arises when dealing with curves. Though a curve is defined to be a function, and its properties are those of functions, we shall also refer to the point set representing the graph of the function as "the curve". Thus, a curve is a continuous function as well as a compact, connected set of points. When the topological properties of a curve are being discussed, the curve will sometimes be denoted by C, without regard to the parameterization from which the curve arose.

For a parameterized curve C defined by (7.5), the point $z(a)$ is called the *initial point* of C and $z(b)$ the *terminal point* of C. If the initial and terminal points coincide, i.e., $z(a) = z(b)$, then C is said to be a *closed curve.* If $z(t_1) \neq z(t_2)$ when $t_1 \neq t_2$, so that C does not intersect itself, the curve is said to be *simple.* A closed curve $C : z(t)$, $t \in [a,b]$, that is simple in the interval (a,b) with the possible exception that $z(a) = z(b)$ is said to be a *simple closed curve* or *Jordan curve.*

Every simple closed curve cuts the plane into two separate domains. In other words, we say that every simple closed curve has an interior (inside) and an outside (exterior). We warn the reader that Jordan curve can be more complicated than Figure 7.1. More formally, we have

Jordan Curve Theorem. *If C is a simple closed (Jordan) curve, then the complement of C consists of two disjoint domains, one bounded domain, and the other an unbounded domain each of which has C as its boundary.*

This geometrically intuitive theorem is remarkably difficult to prove. The reader unwilling to accept the theorem on faith can find a proof in Newman [Ne]. However, given a drawing of some particular simple closed curve, it is usually easy to distinguish the inside from the outside.

A domain D is *simply connected* if each simple closed curve contained in D contains only points of D inside.

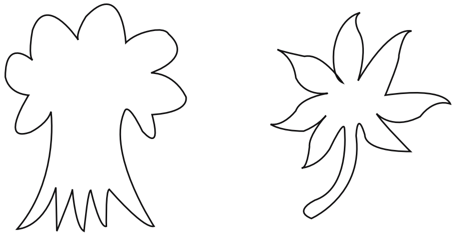


Figure 7.1. Illustration for Jordan curve theorem

For instance, consider the punctured unit disk $D = \{z : 0 < |z| < 1\}$. Then D is a domain but is not simply connected. Some curves, such as C_2, C_3 and C_4 in Figure 7.2 contain only points on D, but C_1 contains $z = 0$ and $z = 0$ does not belong to D. We have the following heuristic interpretations.

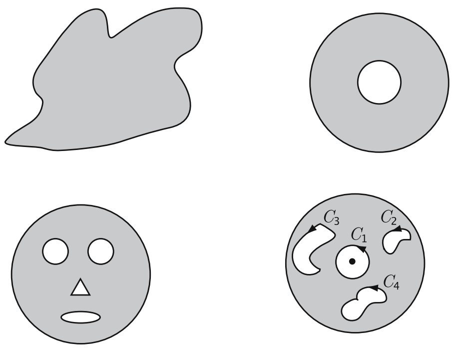


Figure 7.2. Illustration for multiply connected domains

Topologically, a simply connected domain can be continuously shrunk to a point. Note that the punctured unit disk $D = \{z : 0 < |z| < 1\}$ can be shrunk to an arbitrarily small domain, but not to a point in D. Geometrically, a "simply connected domain" has "no holes" inside, for if a simple closed curve should surround a hole, then the curve could not be shrunk beyond the hole. Here again, removal of a single point from a domain is akin to punching a hole in it. A domain that is not simply connected is said to be *multiply connected*.Open disks, open rectangles and star shaped domains are simple examples of simply connected domains. Punctured disks, punctured rectangles, and the punctured plane all have one "hole" and hence are not simply connected. The domain in Figure 7.3 has three holes and hence is not simply connected.

Remark 7.2. In discussing simply connected domains, we will confine ourselves to the finite complex plane. Consequently, the exterior of a circle is not simply connected, since the domain is prevented from being shrunk to a

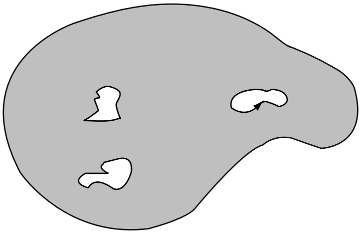


Figure 7.3. Multiply connected domain with three holes

point by the circle from one end and by the point at ∞ from the other. In the extended plane, the exterior of a circle is simply connected because it can be "shrunk" to the point at ∞. ●

Remark 7.3. While we have required analytic functions to be single-valued, it is of interest to discuss a slightly more general concept than analytic, which incorporates multiple-valued functions. We are tempted to say that $\log z$ is analytic in the punctured plane because for each value $z_0 \neq 0$ a branch of $\log z$ may be found in which $\log z$ is analytic at z_0. As we shall see in Chapter 13, this property of the logarithm will enable us to call the multiple-valued function $\log z$ *regular* in the punctured plane. Note that a branch cut for $\log z$ transforms the multiple-valued function into a single-valued function, and also transforms a multiply connected domain (the punctured plane) into a simply connected domain. After the term "regular" is carefully defined in Chapter 13, we shall prove that a function regular in a simply connected domain must also be single-valued (hence analytic) there. This is known as the *Monodromy Theorem.* ●

The boundary C of a domain is said to have *positive orientation*, or to be traversed in the positive sense if a person walking on C always has the domain to his left. The boundary $|z-a| = R$ of the disk $|z-a| < R$ has positive orientation if traversed in a counterclockwise direction and negative orientation if traversed in a clockwise direction. A word of caution: Don't equate positive with counterclockwise. For the annulus in Figure 7.4, the positive direction along the outer circle is counterclockwise, while along the inner circle it is clockwise.

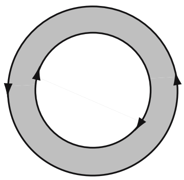


Figure 7.4. An annulus region

However, if a simple closed curve is given without reference to a region, it will be assumed that the domain is inside so that the positive orientation will be counterclockwise.

Remark 7.4. To unquestioningly accept the idea of counterclockwise is to be deluded by the term "simple" closed curve. There are examples of simple closed curves that are not "simple" in the intuitive sense of the word, which occupy almost an entire square. For such a curve, the reader could spend a lifetime tracking down the counterclockwise direction. Although our definition of orientation is more intuitive than rigorous, it will be adequate for all curves encountered in this text. ●

Suppose that

$$\begin{array}{lll} C_1: z_1(t) = e^{it} = \cos t + i \sin t & (0 \leq t \leq 2\pi), \\ C_2: z_2(t) = e^{-it} = \cos t - i \sin t & (0 \leq t \leq 2\pi), \\ C_3: z_3(t) = -e^{it} = -\cos t - i \sin t & (0 \leq t \leq 2\pi), \\ C_4: z_4(t) = -e^{-it} = -\cos t + i \sin t & (0 \leq t \leq 2\pi). \end{array}$$

All four of these simple closed curves traverse the unit circle. They differ from one another either in initial point or in orientation. The curves C_1 and C_2 have initial point $(1, 0)$, whereas C_3 and C_4 have initial point $(-1, 0)$. The curves C_1 and C_3 have positive orientation, and the curves C_2 and C_4 have negative orientation (see Figure 7.5).

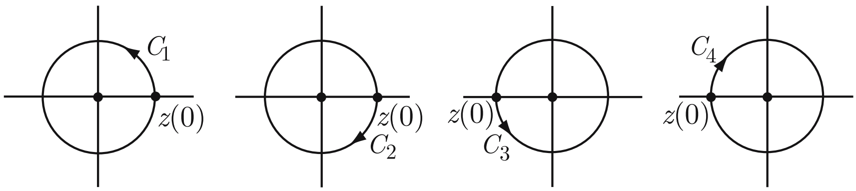


Figure 7.5. Oriented curves

A curve $z(t) = x(t) + iy(t)$, $t \in [a, b]$, having a continuous derivative (i.e., $z(t)$ and $z'(t)$ are continuous on $[a, b]$) is said to be *smooth* or continuously differentiable on $[a, b]$. Of course, by the derivatives at the end points a, b, we mean the appropriate one sided derivatives $z'(a+)$ and $z'(b-)$. For example,

$$z'(a+) = \lim_{t \to a^+} \frac{z(t) - z(a)}{t - a}.$$

A curve γ that is not smooth consists of a finite sequence of smooth curves, $\gamma_1, \gamma_2, \ldots, \gamma_n$ joined together end-to-end. In other words, by a curve we mean a *continuous piecewise smooth curve* defined on a closed interval.

Suppose that $f(x) = u(x) + iv(x)$ is a complex-valued continuous function defined on $[a,b]$. As in the construction of Riemann integral of a real-valued function over $[a,b]$, we consider a partition

$$P:\ a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

and form the corresponding Riemann sum:

$$\sum_{k=1}^{n} f(x_k^*)(x_k - x_{k-1}) = \sum_{k=1}^{n} u(x_k^*)(x_k - x_{k-1}) + i\sum_{k=1}^{n} v(x_k^*)(x_k - x_{k-1})$$

where x_k^* is a point in $[x_{k-1}, x_k]$. As u and v are real-valued continuous on $[a,b]$, the Riemann sum on the right converges to

$$\int_a^b u(x)\,dx + i\int_a^b v(x)\,dx$$

which leads us to define the integration of a complex-valued continuous function of a real variable:

$$\int_a^b f(x)\,dx = \int_a^b u(x)\,dx + i\int_a^b v(x)\,dx.$$

If $f(x)$ is piecewise continuous on $[a,b]$, then apply the above result to each subintervals (a_{k-1}, a_k) $(1 \le k \le m)$ on which $f(x)$ is continuous, and define

$$\int_a^b f(x)\,dx = \sum_{k=1}^{m} \int_{a_{k-1}}^{a_k} u(x)\,dx + i\sum_{k=1}^{m} \int_{a_{k-1}}^{a_k} v(x)\,dx.$$

Thus, (7.1) continues to hold if c_1, c_2 are complex constants and f, g are piecewise continuous complex-valued function defined on $[a,b]$.

More generally, if $f(z)$ is a complex-valued continuous function defined on a smooth curve

$$C:\ z(t) = x(t) + iy(t), \quad t \in [a,b],$$

then it follows that $t \mapsto f(z(t))$ is a continuous function from $[a,b]$ into $\mathbb{C}$. Consequently, $t \mapsto f(z(t))$ is continuous for $a \le t \le b$. We wish to prove that the integral of $f(z)$ on C is given by

$$\int_C f(z)\,dz = \int_a^b f(z(t))z'(t)\,dt. \tag{7.6}$$

We call the right-hand side of the last equation as a "pullback" of the left-hand side of the equation to the interval $[a,b]$. Let us now first define this integral as a limit of sums, analogous to the definition of the Riemann integral. An advantage of (7.6) is that it enables us to use familiar properties of the Riemann integral.

Let P be a partition of $[a,b]$ and $z_k^* = z(t_k^*)$ denotes the point on the subarc with end points $z(t_{k-1})$ and $z(t_k)$. The Riemann sum approximating $\int_C f(z)\,dz$ corresponding to the partition P is given by

$$S_n = \sum_{k=1}^{n} f(z(t_k^*))(z(t_k) - z(t_{k-1})).$$

We have

$$\begin{aligned}
S_n &= \sum_{k=1}^{n}[u(x(t_k^*), y(t_k^*)) + iv(x(t_k^*), y(t_k^*))] \\
&\qquad\qquad [x(t_k) + iy(t_k) - (x(t_{k-1}) + iy(t_{k-1}))] \\
&= \sum_{k=1}^{n} u(x(t_k^*), y(t_k^*))(x(t_k) - x(t_{k-1})) - v(x(t_k^*), y(t_k^*))(y(t_k) - y(t_{k-1})) \\
&\quad + i[u(x(t_k^*), y(t_k^*))(y(t_k) - y(t_{k-1})) + v(x(t_k^*), y(t_k^*))(x(t_k) - x(t_{k-1}))].
\end{aligned}$$

Interpreting each sum on the right as a Riemann sum over the interval $[a,b]$, we have the complex line integral (or contour integral) of f along C as follows:

$$\begin{aligned}
\int_C f(z)\,dz &= \lim_{|P|\to 0} S_n \\
&= \int_a^b u(z)x'(t)\,dt - \int_a^b v(z)y'(t)\,dt \\
&\qquad\qquad + i\int_a^b u(z)y'(t)\,dt + i\int_a^b v(z)x'(t)\,dt \\
&= \int_C f(z(t))\gamma'(t)\,dt,
\end{aligned}$$

where $|P|$ denotes the maximum of the length of the subintervals. Thus, we have actually proved (7.6). We may conveniently write the last expression as

$$\int_C f(z)\,dz = \int_a^b [ux' - vy']\,dt + i\int_a^b [uy' + vx']\,dt. \tag{7.7}$$

Thus, just as a complex function may be expressed in terms of real-valued functions, so may a complex integral clearly be expressed in terms of real-valued integrals. We formulate the above discussion as

Theorem 7.5. *Suppose that $f(z) = u(x,y) + iv(x,y)$ is continuous on a parameterized smooth curve $C:\ z(t) = x(t) + iy(t)$, $t \in [a,b]$. Then*

$$\int_C f(z)\,dz = \int_C u\,dx - v\,dy + i\int_C u\,dy + v\,dx. \tag{7.8}$$

Observe that the right side of (7.8) would be obtained by the formal substitution

$$f = u + iv, \quad dz = dx + i\,dy$$

into the left side of (7.8). Either equation (7.7) or (7.8) could have been taken as the definition of the complex integral, instead of (7.6).

Also, we observe that the integrand on the right side of (7.6) would be obtained by the formal substitution

$$z = z(t), \quad dz = z'(t)\,dt$$

into the left side of (7.6). Moreover, in the special case that $z(t) = t$, the curve is a real interval and (7.6) reduces to an integral of the form (7.3).

For a general piecewise smooth curve C, the derivative $z'(t)$ need not be continuous but is piecewise continuous so that

$$t \mapsto f(z(t))z'(t)$$

is piecewise continuous. In this case we evaluate the integral as a finite sums of integrals of continuous functions. The above discussion leads to

Definition 7.6. *Let C be a piecewise smooth curve on $[a,b]$ and f a continuous function on the graph/trace of C. The contour integral of f along C is defined to be*

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

Sometimes the notation $\int_\gamma f(z)\,d\gamma$ or $\int_\gamma f\,d\gamma$ is used when γ is a piecewise smooth curve.

An expression of the form $\int_C P(x,y)\,dx + Q(x,y)\,dy$ is called a real *line integral.* From (7.7), we see that the complex (line) integral may be expressed in terms of two real line integrals. We give here an example to illustrate different methods for computing a complex integral.

Example 7.7. Consider the problem of evaluating $I = \int_\gamma z^2\,dz$, where

(i) γ is an arc of a circle centered at the origin
(ii) γ is the union of the horizontal segment from 0 to 1 and the vertical segment from 1 to $1 + 2i$
(iii) γ is the line segment from 0 to $1 + 2i$
(iv) γ is the contour parameterized by $\gamma :\ z(t) = t^2 + it$ $(0 \leq t \leq 1)$.

Let $f(z) = z^2$. In the first case we may write γ in the form

$$\gamma(t) = re^{it}, \quad a \leq t \leq b.$$

Then $\gamma'(t) = ire^{it}$ and $f(\gamma(t))\gamma'(t) = ir^3e^{3it}$ so that

$$I = \int_a^b f(\gamma(t))\gamma'(t)\,dt = r^3 \frac{e^{3ib} - e^{3ia}}{3}.$$

In particular if γ is a closed circle, then $I = 0$, since $b = a + 2k\pi$ for some integer k. In the second case, we may write γ as

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

Therefore, we have

$$I = \int_0^1 f(\gamma(t))\gamma'(t)\,dt + \int_1^3 f(\gamma(t))\gamma'(t)\,dt = \int_0^1 t^2\,dt + \int_1^3 [1 + (t-1)i]^2 i\,dt$$

and, it is a simple exercise to see that $I = -(11 + 2i)/3$.

In the third case, the path γ is given by $\gamma(t) = t(1 + 2i)$, $0 \le t \le 1$. Therefore the integral is

$$\int_0^1 [t^2(1+2i)^2](1+2i)\,dt = (1+2i)^3 \left.\frac{t^3}{3}\right|_0^1 = -\frac{11+2i}{3}.$$

In the final case, according to (7.6), we can easily see that

$$\int_\gamma z^2\,dz = -\frac{2}{3} + \frac{2}{3}i. \qquad \bullet$$

Remark 7.8. At first glance, it appears that (7.8) serves no purpose other than to introduce a cumbersome method for evaluating the complex integral. We will rarely use (7.8) to compute integral directly. However, it will enable us to formulate theorems about the complex integral from theorems about real line integrals. This will lead to a method for evaluating the complex integral that is far simpler than (7.6). •

Example 7.9. Consider the curve γ given by

$$z(t) = \begin{cases} t(1 + it\sin(1/t)) & \text{if } t \ne 0 \\ 0 & \text{if } t = 0. \end{cases}$$

Then

$$z'(t) = \begin{cases} 1 + i(2t\sin(1/t) - \cos(1/t)) & \text{if } t \ne 0 \\ 1 & \text{if } t = 0. \end{cases}$$

Note that $z'(t)$ is discontinuous at 0 and neither the left nor the right limit of $z'(t)$ exists at 0. So, $z'(t)$ is not piecewise continuous, for example on $[-\pi, \pi]$. Consequently, the restriction of the curve γ to $[-\pi, \pi]$ is not smooth. •

Remark 7.10. The value of the real integral $\int_a^b f(x)\,dx$ depends on the function $f(x)$ and the end points of the interval $[a, b]$. The value of the complex integral $\int_C f(z)\,dz$ may depend on the function $f(z)$ and *all* the points on the curve C, not just the end points of C. •

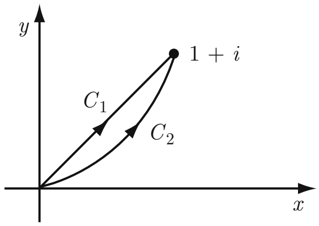


Figure 7.6. Graph of curves C_1 and C_2

Example 7.11. We wish to find $\int_C |z|^2\,dz$ along the curves

(a) $C = C_1 : z_1(t) = t + it \quad (0 \le t \le 1)$,
(b) $C = C_2 : z_2(t) = t^2 + it \quad (0 \le t \le 1)$.

Clearly, C_1 and C_2 are smooth and $f(z) = |z|^2$ is continuous on $\mathbb{C}$. According to (7.6),

$$\int_{C_1} |z|^2\,dz = \int_0^1 |t + it|^2(1+i)\,dt = (1+i)\int_0^1 2t^2\,dt = \frac{2}{3} + \frac{2}{3}i$$

and similarly,

$$\int_{C_2} |z|^2\,dz = \int_0^1 |t^2 + it|^2(2t+i)\,dt = \frac{5}{6} + \frac{8}{15}i.$$

Despite the fact that the straight line C_1 and the parabola C_2 have the same initial and terminal points (see Figure 7.6), we have

$$\int_{C_1} |z|^2\,dz \ne \int_{C_2} |z|^2\,dz.$$

Note, however, that

$$\int_{C_1} z^2\,dz = \int_0^1 (t+it)^2(1+i)\,dt = 2i(1+i)\frac{1}{3} = \int_{C_1} z^2\,dz.$$

It is no coincidence that

$$\int_{C_1} z^2\,dz = \int_{C_2} z^2\,dz.$$

It will later be shown that the value of $\int_C z^2\,dz$ depends only on the initial and terminal points of the smooth curve C. Our goal in this chapter is to characterize the class of functions for which the integral is independent of path, i.e., functions for which

$$\int_{C_1} f(z)\,dz = \int_{C_2} f(z)\,dz$$

along any smooth curves C_1 and C_2 having the same initial and terminal points. ●

Questions 7.12.

1. Is a single "point" a curve?
2. Is a simple closed curve a simple curve?
3. In our definition of curve, would it have made any difference had the parameter t been restricted to the interval [0,1]?
4. What is the relationship between simply connected and connected?
5. Can a domain and its complement both be simply connected?
6. Can a domain have finitely many holes?
7. Can a domain have infinitely many holes?
8. When a point is removed from a simply connected domain, is the new domain simply connected?
9. Why was it important for the derivative of a curve to be continuous?
10. What is the relationship between an integral being independent of path and an integral around a closed curve being zero?
11. Why is every curve compact and connected?
12. Let D be the complement in $\mathbb{C}$ of the real axis, i.e., $D = \mathbb{C}\setminus\mathbb{R}$. Is D simply connected?
13. Let D be the complement in $\mathbb{C}$ of the nonnegative real axis, i.e., $D = \mathbb{C}\setminus\{x \in \mathbb{R} : x \geq 0\}$. Is D simply connected?

Exercises 7.13.

1. Describe the curve $z(t) = a\cos t + ib\sin t$ $(-\pi \leq t \leq \pi)$, where a and b are positive real numbers.
2. Describe the curve $z(t) = t^3 + it^2$ $(-1 \leq t \leq 1)$. Is it a "smooth curve"?
3. Describe the curve
$$z(t) = \frac{1-t^2}{1+t^2} + i\frac{2t}{1+t^2} \quad (-R \leq t \leq R).$$
What happens as $R \to \infty$?
4. Plot the given curves
(i) $z(t) = \begin{cases} t & \text{if } -3 \leq t \leq -1 \\ e^{i\pi(1-t)/2} & \text{if } -1 \leq t \leq 1 \\ t & \text{if } 1 \leq t \leq 3. \end{cases}$
(ii) $z(t) = \begin{cases} t(1+i) & \text{if } 0 \leq t \leq 1 \\ 3+i-2t & \text{if } 1 \leq t \leq 2 \\ (-1+i)(3-t) & \text{if } 2 \leq t \leq 3. \end{cases}$
5. Find a parameterized curve tracing out the following loci:
   (a) The line segment from $z = i$ to $z = 1 - i$
   (b) The line segment from $z = 1$ to $z = 2 + 3i$
   (c) The square whose vertices are $\pm 1 \pm i$, traversed in the positive sense, with initial point $-1 - i$
   (d) The part of the circle $|z - 1| = 2$ in the right half-plane.
6. Find a parameterized curve for the parabola $y = 2x^2 - 3$ that has initial point $z = -1 - i$ and terminal point $z = 2 + 5i$.

7. Parameterize the following simple closed curves in polar coordinates.

$$\text{(a)}\quad x^2+y^2=4 \qquad \text{(b)}\quad 4x^2+y^2=1 \qquad \text{(c)}\quad x^2+(y+1)^2=9.$$

8. Find $\int_C \overline{z}\,dz$ along the following curves.
   (a) $z(t)=e^{it}$ $(-\pi\le t\le\pi)$
   (b) $z(t)=e^{2it}$ $(-\pi\le t\le\pi)$
   (c) $z(t)=e^{it}-1$ $(-\pi\le t\le\pi)$
   (d) $z(t)=t+it$ $(0\le t\le 2)$
   (e) $z(t)=5e^{it}+3$ $(0\le t\le\pi)$
   (f) $z(t)=1-t+it$ $(0\le t\le 1)$
   (g) $z(t)=1+it$ $(0\le t\le 1)$
   (h) $z(t)=1+i-t$ $(0\le t\le 1)$.
9. Along the curve $C:\ z(t)=e^{it}\quad(-\pi\le t\le\pi)$, evaluate $\int_C f(z)\,dz$ for
   (a) $f(z)=z^2$ (b) $f(z)=\dfrac{1}{z}$
   (c) $f(z)=\dfrac{1}{z^2}$ (d) $f(z)=2\overline{z}-i\left(z+\dfrac{1}{z}\right)$.

7.2 Parameterizations

Suppose $C:\ z(t)$ is a smooth curve defined on the interval $[a,b]$. Breaking the interval into two subintervals $[a,c]$ and $[c,b]$, we obtain two curves C_1 and C_2 from $z(t)$ by restricting the parameter t to the intervals $[a,c]$ and $[c,b]$, respectively. For any function $f(z)$ continuous on C,

$$\begin{aligned}\int_C f(z)\,dz &= \int_a^b f(z(t))z'(t)\,dt\\ &= \int_a^c f(z(t))z'(t)\,dt + \int_c^b f(z(t))z'(t)\,dt\\ &= \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz.\end{aligned}$$

Similarly, the curve C can be expressed as the "sum" of n curves with

$$\begin{aligned}\int_C f(z)\,dz &= \int_{C_1+\cdots+C_n} f(z)\,dz\\ &= \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz + \cdots + \int_{C_n} f(z)\,dz.\end{aligned} \tag{7.9}$$

Remark 7.14. By the sum of two curves, we mean the curve formed by joining the initial point of one curve to the terminal point of the other; this is not to be confused with termwise addition of functions defined on the same set. ●

A function is *sectionally continuous* on an interval if it has at most a finite number of discontinuities, with right- and left-hand limits at each point in the interval. More precisely, f is sectionally continuous on $[a, b]$ if

(i) f is continuous at all but finitely many points on (a, b).
(ii) At any point c in (a, b) where f fails to be continuous, both left limit $\lim_{t\to c^-} f(t)$ and the right limit $\lim_{t\to c^+} f(t)$ exist and are finite.
(iii) At the end points, the right limit $\lim_{t\to a^+} f(t)$ and the left limit $\lim_{t\to b^-} f(t)$ exist and are finite.

A curve having a sectionally continuous derivative is called a *contour*. In other words, *piecewise smooth curve* is called a *contour*. Recall that a path $C : z(t), t \in [a, b]$ is said to be *piecewise smooth* if there exists a partition $P : a = t_0 < t_1 < \cdots < t_n = b$ of $[a, b]$ such that the restriction of C to each of the subintervals $[t_{k-1}, t_k]$, $k = 1, 2, \ldots, n$, is a smooth curve.

Since every contour C may be expressed as the sum of a finite number of smooth curves, $C_1 + C_2 + \cdots + C_n$, the integral of a continuous function along a contour is defined by (7.9).

Example 7.15. The curve

$$C : z(t) = t + i|t|, \quad t \in [-1, 1],$$

is a piecewise smooth but not a smooth curve. It is easy to see that the derivative at the origin fails to exist. The restriction of C to $[-1, 0]$ and to $[0, 1]$ is clearly seen to be smooth. Hence, C is referred to as a contour. Note also that C is simple but not closed. How about the curve described by $z(t) = |t| + it$, $t \in [-1, 1]$? How about the curve described by $z(t) = |t^3| + it^3$ on the interval $t \in [-1, 1]$? ●

Define $C = C_1 + C_2 + C_3 + C_4$, where C_j's are the line segments given by $C_1 = [0, 1]$, $C_2 = [1, 1 + i]$, $C_3 = [1 + i, i]$, $C_4 = [i, 1]$; see Figure 7.8. Then C describes the boundary of a square. Note that the curve C is piecewise smooth, simple and closed.

Example 7.16. We find the value of the integral $\int_C z\,dz$ along the contour

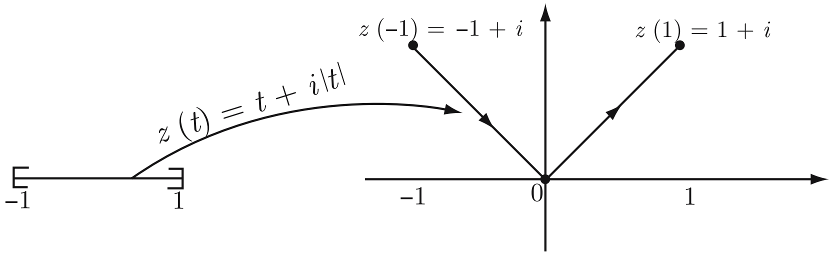


Figure 7.7. The piecewise smooth curve $C : z(t) = t + i|t|, \; t \in [-1, 1]$

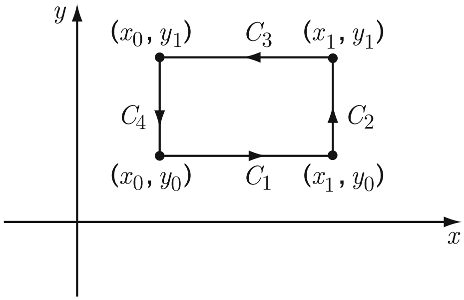


Figure 7.8. The curve $C = C_1 + C_2 + C_3 + C_4$

$$C: \; z(t) = \begin{cases} 2t & \text{if } 0 \le t \le 1, \\ 2 + i(t-1) & \text{if } 1 \le t \le 2. \end{cases}$$

Defining curves C_1 and C_2 by restricting the parameter t of C to the intervals $[0, 1]$ and $[1, 2]$, respectively (see Figure 7.9), we have

$$\begin{aligned} \int_C z\,dz &= \int_{C_1} z\,dz + \int_{C_2} z\,dz \\ &= \int_0^1 2t(2\,dt) + \int_1^2 \{2 + i(t-1)\}(i\,dt) \\ &= \int_0^1 4t\,dt - \int_1^2 (t-1)\,dt + \int_1^2 2i\,dt \\ &= 2 - \frac{1}{2} + 2i = \frac{3}{2} + 2i. \end{aligned}$$

●

Recall, from elementary calculus, that the (arc) length L of a smooth curve in the plane defined parametrically by the equations

$$x = \phi(t), \quad y = \psi(t) \quad (a \le t \le b)$$

is given by

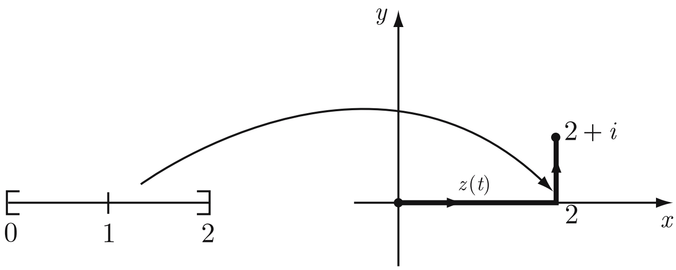


Figure 7.9.

$$L = \int_a^b \sqrt{(\phi'(t))^2 + (\psi'(t))^2}\, dt = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.$$

Note that the integrand on the right integral is recognized as ds/dt, where s is the arc measured from the point $z(a)$ of C. Using the parameterization

$$C:\ z(t) = x(t) + iy(t) \quad (a \le t \le b)$$

for C, a smooth curve (or contour) in the plane, the length of C is given by

$$L = \int_a^b |z'(t)|\, dt = \int_a^b \left|\frac{dx}{dt} + i\frac{dy}{dt}\right| dt. \tag{7.10}$$

In the special case that the curve is a line segment from z_0 to z_1, parameterized by

$$z(t) = tz_0 + (1-t)z_1 \quad (0 \le t \le 1),$$

we have $z'(t) = z_1 - z_0$. Hence, as expected,

$$L = \int_0^1 |z'(t)|\, dt = \int_0^1 |z_1 - z_0|\, dt = |z_1 - z_0|.$$

When z is on C, the symbol $|dz| = |z'(t)|\, dt$ so that (7.10) may also be written as

$$L = \int_C |dz| = \int_C ds, \tag{7.11}$$

it being understood that C is parameterized by $z(t)$. This observation partly explains why $\int_C f(z)\,|dz|$ defines an integral of f along C with respect to arc length:

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

The arc length integrals of this type play significant roles in certain areas of mathematics and physics.

Example 7.17. Let us now evaluate $\int_C z^{-n}\,|dz|$ where $C = z(t) = re^{it}$ ($r > 0,\ 0 \le t \le 2\pi$ and $n \in \mathbb{Z}$). Set $f(z) = 1/z^n$. Note that C is smooth and

$$z'(t) = ire^{it}, \quad f(z(t)) = \frac{e^{-int}}{r^n}.$$

Therefore,

$$\int_C \frac{1}{z^n}\,|dz| = \int_0^{2\pi} \frac{r\, dt}{r^n e^{int}} = \frac{1}{r^{n-1}} \int_0^{2\pi} e^{-int}\, dt = \begin{cases} 0 & \text{if } n \ne 0 \\ 2\pi r & \text{if } n = 0. \end{cases}$$ ●

Remark 7.18. It is meaningful to talk about the length of an arbitrary curve. However, if the curve is not a contour, then the length may not be finite. Consider an arbitrary curve $z(t)$, with $a \le t \le b$. Define

$$V(P) = \sum_{k=1}^{n} |z(t_k) - z(t_{k-1})|, \tag{7.12}$$

where $a = t_0 < t_1 < \cdots < t_n = b$ is a partition P. By the triangle inequality, $V(P)$ increases monotonically as the subintervals are further subdivided into smaller subintervals. The length of this curve can be defined as the least upper bound of all sums of the form (7.12), that is,

$$\sup_P \sum_{k=1}^{n} |z(t_k) - z(t_{k-1})|.$$

If the length is finite, the curve is said to be *rectifiable*. The reader should verify that every contour is rectifiable and that, in the case of a contour, this definition agrees with (7.11). In Exercise 7.29(1), an example of a nonrectifiable curve is given. ●

What follows is the complex analog to a well-known real variable theorem.

Theorem 7.19. (M-L Inequality) *Suppose $f(z)$ is continuous on a contour C having length L, with $|f(z)| \le M$ on C. Then*

$$\left|\int_C f(z)\,dz\right| \le \int_C |f(z)|\,|dz| \le M\int_C |dz| = ML.$$

Proof. Since C is parameterized by $z(t)$ on the interval $[a, b]$, we have

$$\begin{aligned}\left|\int_C f(z)\,dz\right| &= \left|\int_a^b f(z(t))z'(t)\,dt\right| \\ &\le \int_a^b |f(z(t))|\,|z'(t)|\,dt = \int_C |f(z)|\,|dz| \\ &\le M\int_a^b |z'(t)|\,dt = ML \quad (\text{since } |f(z)| \le M \text{ on } C)\end{aligned}$$

and the conclusion follows. ∎

For example, we can use the M-L Inequality to find an upper bound for $|\int_C (z^2+10)^{-1}\,dz|$, where C is the circle $C:\ z(t) = 2e^{it}$ $(-\pi \le t \le \pi)$. In fact, for $z \in C$, $|z^2+10| \ge 10 - |z|^2 = 10 - |z(t)|^2 \ge 10 - 4 = 6$, and so, we have

$$\left|\int_C \frac{dz}{z^2+10}\right| \le \int_C \frac{|dz|}{|z^2+10|} \le \frac{1}{6}\int_C |dz| = \frac{2\pi}{3}.$$

Similarly, we easily see that

$$\left|\int_C \frac{e^z}{z+1}\,dz\right| \leq 4\pi e^2.$$

Strictly speaking, a curve C is associated with a definite parametric form $C:\, z = z(t)$, $a \leq t \leq b$ and so, the length of a curve is defined in terms of its parameter; but it is geometrically evident that, for simple curves, the length is independent of the parameterization. For instance, the curves

$$C_1:\, z_1(t) = e^{it} \quad (0 \leq t \leq \pi), \quad \text{and} \quad C_2:\, z_2(t) = e^{2it} \quad (0 \leq t \leq \pi/2)$$

both traverse the upper half of the unit circle. Moreover, by (7.11), we have

$$\int_{C_1} |dz_1| = \int_0^{\pi} |z_1'(t)|\,dt = \int_0^{\pi} dt = \pi,$$

and

$$\int_{C_2} |dz_2| = \int_0^{\pi/2} |z_2'(t)|\,dt = \int_0^{\pi/2} 2\,dt = \pi.$$

The contours C_1 and C_2, although different in formal sense because they arise from different parameterizations, have the same length.

Remark 7.20. The curves e^{it} $(0 \leq t \leq 2\pi)$ and e^{2it} $(0 \leq t \leq 2\pi)$ both represent the set of points on the unit circle. The length of the first curve is 2π and that of second is 4π. Note, however, that the second curve is not simple because it traverses the unit circle twice. ●

The next theorem gives general criteria for changing parameters without affecting arc length.

Theorem 7.21. *Let $C:\, z(t) = x(t) + iy(t)$, $a \leq t \leq b$, be a contour. Suppose $t = \phi(s)$ with $a = \phi(c)$, $b = \phi(d)$, and $\phi'(s) > 0$, so that t increases with s. If $\phi'(s)$ is sectionally continuous on the interval $[c, d]$, then the length of C is given by*

$$L = \int_c^d |z'(\phi(s))|\phi'(s)\,ds.$$

Proof. By (7.10) and the chain rule,

$$\begin{aligned} L &= \int_a^b |z'(t)|\,dt = \int_a^b |x'(t) + iy'(t)|\,dt \\ &= \int_a^b |x'(\phi(s)) + iy'(\phi(s))|\,d\phi(s) \\ &= \int_a^b |z'(\phi(s))|\phi'(s)\,ds. \end{aligned}$$

■

Let C be an arc $z(t)$, $-2 \leq t \leq 1$ and let C_1 be the arc $z(s) = \gamma(3s-5)$, $1 \leq s \leq 2$. Clearly, C and C_1 have the same initial and the same end point, and the same trajectory. It is easy to see that

$$\int_C f(z)\,dz = \int_{C_1} f(z)\,dz \tag{7.13}$$

holds for every continuous function in D which contains C_1 and C_2. Indeed, if we set $t = \phi(s) = 3s-5$, then $\phi(1) = -2$, $\phi(2) = 1$ and $\phi'(s) = 3$. Hence

$$\int_{C_1} f(z)\,dz = \int_1^2 f(\gamma(3s-5))3\gamma'(3s-5)\,ds.$$

By the substitution $t = \phi(s)$, the change of variable leads to $\int_1^2 f(z(t))z'(t)\,dt$ which is nothing but the right-hand side of (7.13). This continues to hold for a broader class of functions as we see next.

Suppose that $f(z)$ is continuous on a contour $C:\, z(t)$, $a \leq t \leq b$, and that $t = \phi(s)$ satisfies the conditions of Theorem 7.21. Then the chain rule may be applied to obtain

$$\begin{aligned} \int_C f(z)\,dz &= \int_a^b f(z(t))z'(t)\,dt \qquad (7.14) \\ &= \int_c^d f(z(\phi(s)))z'(\phi(s))\phi'(s)\,ds. \end{aligned}$$

Since $z'(\phi(s))\phi'(s)$ is the derivative with respect to s of $z(\phi(s))$, the curve C could have been parameterized by $z(\phi(s))$, $c \leq s \leq d$, without affecting the value of the integral.

The contour

$$-C:\, z(-t) = x(-t) + iy(-t) \quad (-b \leq t \leq -a)$$

represents the same curve, traversed in the opposite direction, as

$$C:\, z(t) = x(t) + iy(t) \quad (a \leq t \leq b).$$

We have

$$\int_{-C} f(z)\,dz = \int_{-b}^{-a} f(z(-t))z'(-t)(-1)\,dt$$

and, upon making the substitution $s = -t$,

$$\begin{aligned} \int_{-C} f(z)\,dz &= \int_b^a f(z(s))z'(s)\,ds \qquad (7.15) \\ &= -\int_a^b f(z(s))z'(s)\,ds \\ &= -\int_C f(z)\,dz. \end{aligned}$$

Remark 7.22. Loosely speaking, equations (7.14) and (7.15) say that the value of the integral along a simple contour C, viewed as a point set in the plane, depends on the parameterization of the contour only with regard to orientation. ●

Remark 7.23. Since integrating around a circle is such a common occurrence, we introduce the notation $\int_{|z-z_0|=r} f(z)\,dz$, which will be interpreted as the integral of $f(z)$ around the contour consisting of the circle $|z-z_0|=r$ oriented in the positive sense. ●

Examples 7.24. Let us evaluate $\int_C |z|^n\,dz$ ($n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$) along the straight line C joining the origin to the point $1+i$. We parameterize the line by

$$C:\ z(t) = t + it \quad (0 \le t \le 1). \tag{7.16}$$

Then $z'(t) = 1 + i$, and

$$\int_C |z|^n\,dz = \int_0^1 |t+it|^n(1+i)\,dt = 2^{n/2}(1+i)\int_0^1 t^n\,dt = \frac{2^{n/2}(1+i)}{n+1}.$$

The parameterization (7.16) was chosen because it was the most natural. We could have parameterized C by

$$C:\ z(t) = x(t) + ix(t) \quad (a \le t \le b),$$

where $x(a) = 0$, $x(b) = 1$, and $x'(t) > 0$. Then $z'(t) = (1+i)x'(t)$, and

$$\begin{aligned}\int_C |z|^n\,dz &= \int_a^b |x(t)+ix(t)|^n(1+i)x'(t)\,dt\\ &= 2^{n/2}(1+i)\int_a^b (x(t))^n x'(t)\,dt\\ &= 2^{n/2}(1+i)\left(\frac{(x(b))^{n+1}-(x(a))^{n+1}}{n+1}\right) = \frac{2^{n/2}(1+i)}{n+1}.\end{aligned}$$

Next, to evaluate

$$\int_{|z|=r} |z|^n\,dz \quad (n \in \mathbb{Z}),$$

we parameterize the specified circle by $z(t) = re^{it}$, $0 \le t \le 2\pi$, so that

$$\int_C |z|^n\,dz = \int_0^{2\pi} |re^{it}|^n ire^{it}\,dt = ir^{n+1}\int_0^{2\pi} e^{it}\,dt = 0.$$ ●

Example 7.25. Let us now consider one more similar integral over a closed contour. Consider $\int_C |z|\,dz$ along the rectangle C having corners $-1, 1, 1+i, -1+i$. This contour is the sum of four smooth curves (straight lines). We have

$$\int_C |z|\,dz = \int_{C_1} |z|\,dz + \int_{C_2} |z|\,dz + \int_{C_3} |z|\,dz + \int_{C_4} |z|\,dz,$$

where

$$\begin{array}{lll} C_1 : & z_1(t) = t & (-1 \le t \le 1), \\ C_2 : & z_2(t) = 1 + it & (0 \le t \le 1), \\ C_3 : & z_3(t) = -t + i & (-1 \le t \le 1), \\ C_4 : & z_4(t) = -1 - it & (-1 \le t \le 0). \end{array}$$

Solving, we obtain

$$\begin{aligned} \int_C |z|\,dz &= \int_{-1}^{1} |t|\,dt + i\int_0^1 \sqrt{1+t^2}\,dt - \int_{-1}^{1} \sqrt{t^2+1}\,dt - i\int_{-1}^{0} \sqrt{1+t^2}\,dt \\ &= \int_{-1}^{1} (|t| - \sqrt{t^2+1})\,dt \\ &= 2\int_0^1 (t - \sqrt{t^2+1})\,dt \\ &= 1 - \sqrt{2} - \ln(\sqrt{2}+1). \end{aligned}$$

●

Remark 7.26. The contour C is not the sum of the four curves C_1, C_2, C_3, C_4 as defined in (7.9), because these curves are not parameterized on four distinct subintervals of the interval on which C is parameterized; but according to (7.14), the parameterization is not critical, so we will adopt a more liberal definition of "sum" that does not require a specific parameterization. In fact, we may even integrate without expressing x and y in terms of common parameter t. Therefore, in the last example, we could just as well have written

$$\int_C |z|\,dz = \int_{-1}^{1} |x|\,dx + \int_0^1 |1+iy| i\,dy + \int_1^{-1} |x+i|\,dx + \int_1^0 |-1+iy| i\,dy.$$

Here, we are actually using x for the parameter on C_1 and C_3, and y for the parameter on C_2 and C_4. ●

Examples 7.27. For each integer n, we have

$$\int_{|z|=r} z^n\,dz = \int_0^{2\pi} (re^{it})^n ire^{it}\,dt = ir^{n+1}\int_0^{2\pi} e^{i(n+1)t}\,dt = \begin{cases} 0 & \text{if } n \ne -1, \\ 2\pi i & \text{if } n = -1. \end{cases}$$

Note that the value of this integral is independent of the radius of the given circle.

Our final example of this section is to compute $I = \int_{|z|=r} x\,dz$. Note that $|z|^2 = z\overline{z} = r^2$ and $x = (z+\overline{z})/2$. Thus, the linearity property gives

$$I = \frac{1}{2}\int_{|z|=r} (z + r^2/z)\,dz = \frac{1}{2}\int_{|z|=r} z\,dz + \frac{r^2}{2}\int_{|z|=r} \frac{dz}{z} = i\pi r^2.$$

Note that the value of the integral in this case depends on the radius of the given circle. Why is this so? Again, as $\overline{z} = r^2/z$, we have

$$\int_{|z|=r} (\overline{z})^n\, dz = r^{2n} \int_{|z|=r} \frac{1}{z^n}\, dz = \begin{cases} 2\pi i r^2 & \text{if } n = 1 \\ 0 & \text{if } n \in \mathbb{Z} \setminus \{1\}. \end{cases}$$ •

Questions 7.28.

1. Can (7.9) be extended to the case of infinitely many curves?
2. How may the following expressions be interpreted?

$$\text{(a)} \int_C |f(z)|\, dz \qquad \text{(b)} \int_C f(z)\, |dz| \qquad \text{(c)} \int_C |f(z)|\, |dz|.$$

3. Can $\int_C f(z)\, dz$ be defined without requiring $f(z)$ to be continuous at all points of C?
4. If $f(z)$ is continuous on a contour C, does $|f(z)|$ necessarily assume a maximum on C?
5. If $f(z)$ is continuous for $|z| < r$ for some $r > 0$ such that $f(0) = 0$, is $\lim_{\delta \to 0} \int_0^{2\pi} f(\delta e^{i\theta})\, d\theta = 0$? Is $\lim_{\delta \to 0} \int_{|z|=\delta} \frac{f(z)}{z}\, dz = 0$?
6. Does the orientation affect the length of a curve?
7. Why is it usually easier to integrate along a circle than along a square?
8. If $|f(z)| \leq 2$ on the circle $|z| = 3$, is $\left| \int_{|z|=3} f(z)\, dz \right| \leq 3$?

Exercises 7.29.

1. Show that the curve C parameterized by

$$z(t) = \begin{cases} t\left(\cos\left(\frac{1}{t}\right) + i \sin\left(\frac{1}{t}\right)\right) & \text{if } 0 < t \leq 1 \\ 0 & \text{if } t = 0, \end{cases}$$

is nonrectifiable.
2. Prove that every (continuous) curve is bounded.
3. Show that

$$\text{(a)} \left| \int_{|z|=1} \frac{dz}{3 + 5z^2} \right| \leq \pi \qquad \text{(b)} \left| \int_{|z|=1} \frac{2z + 1}{5 + z^2}\, dz \right| \leq \frac{3\pi}{2}.$$

4. Find the length of the following contours.
   (a) $z(t) = 3e^{2it} + 2 \qquad (-\pi \leq t \leq \pi)$
   (b) $z(t) = e^t \cos t + i e^t \sin t \qquad (-\pi \leq t \leq \pi)$.
5. Evaluate $\int_C x\, dz$, $\int_C y\, dz$, $\int_C \overline{z}\, dz$ along the following contours:
   (a) The line segment from the origin to $1 + i$
   (b) The line segment from the origin to $1 - i$
   (c) The circle $|z| = 1$
   (d) The curve C consisting of the line segment from 0 to 1 followed by the line segment from 1 to $1 + i$
   (e) The curve C consisting of the line segment from 0 to i followed by the line segment from i to $1 + i$.

6. Evaluate $\int_C z\,dz$, $\int_C |z|\,dz$, $\int_C \overline{z}|z|\,dz$, $\int_C z\,|dz|$, $\int_C |z|\,|dz|$ along the same contours as above. Do the same for the closed contour C consisting of the upper semicircle $|z| = 1$ from 1 to -1, and the line segment $[-1, 1]$.
7. Evaluate $\int_C |z|^2\,dz$, $\int_C \operatorname{Re} z\,|dz|$ and $\int_C \operatorname{Im} z\,|dz|$, where $C:\ z(t) = t^2/3 + it$ for $0 \le t \le 1$.
8. Evaluate $\int_C (az + b\overline{z})\,dz$ where a, b are some nonzero fixed constants and C is the contour given by $C = [0, e^{i\pi/6}] \cup \{e^{i\theta} : \pi/6 \le \theta \le \pi/2\} \cup [e^{i\pi/3}, 0]$. Do the same by replacing $\pi/6$ by α and $\pi/3$ by β, $0 \le \alpha < \beta \le 2\pi$. Do the same for $C:\ z(t) = -t + i(t^2 + 2)$, $0 \le t \le 2$.
9. Evaluate $\int_C (1/z)\,dz$ along the square having corners $\pm 1 \pm i$.
10. Evaluate the following integrals:
    (a) $\int_C e^z\,dz$ along the line segment from the origin to $2 + 2i$
    (b) $\int_C (e^z + z + 1)\,dz$ along the line segment from $-1 + i$ to $1 - i$
    (c) $\int_C \cos z\,dz$ along the line segment from the origin to the point $1 + i$
    (d) $\int_C |z|^2\,dz$ along the square with vertices $0, 1, 1 + i, i$
    (e) $\int_C (x^2 + iy^2)\,dz$ along the line segment from 0 to $1 + i$ followed by the line segment from $1 + i$ to $1 + 2i$.
11. Evaluate $\int_C (z/\overline{z})\,dz$ along the simple closed contour C shown in Figure 7.10.

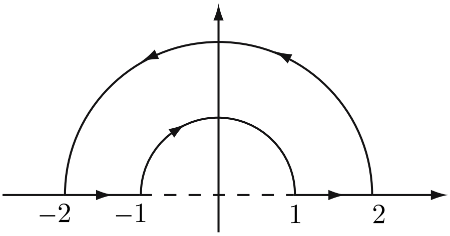


Figure 7.10.

12. Evaluate $\int_C z|z|\,dz$ along the upper semicircle $|z| = R$ from R to $-R$, and the line segment $[-R, R]$.

7.3 Line Integrals

In order to draw a useful analogue with single-variable calculus, we begin by reviewing the (first) fundamental theorem of calculus. Suppose $f(x)$ is continuous on the interval $[a, b]$. The first fundamental theorem of calculus asserts the existence of an antiderivative $F(x)$ for $f(x)$ (i.e., a function F such that $F'(x) = f(x)$ on $[a, b]$) with

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

This theorem relates the behavior of a function on the boundary of a set (two points) to the behavior of an associated function, its derivative, on the whole set (a closed interval). From our earlier discussion, it is clear that the familiar properties of Riemann integral is carried over to the case of complex integrals. For instance, if $f,\, g:\, [a,b] \to \mathbb{C}$ are continuous and if c is a complex constant, then

$$\int_a^b (f(x) + cg(x))\, dx = \int_a^b f(x)\, dx + c\int_a^b g(x)\, dx.$$

The fundamental theorem of calculus is also valid in this setting. More precisely, we have

Theorem 7.30. *If $f:\, [a,b] \to \mathbb{C}$ is continuous and if there exists a function $F(x)$ such that $F'(x) = f(x)$ on $[a,b]$, then*

$$\int_a^b f(x)\, dx = F(x)|_a^b = F(b) - F(a).$$

For instance, if $f_1(t) = 3t^2 - 2it$ and $f_2(t) = e^{2\pi it}$, then the corresponding antiderivatives are $F_1(t) = t^3 - it^2$ and $F_2(t) = e^{2\pi it}/(2\pi i)$ so that

$$\int_0^1 (3t^2 - 2it)\, dt = t^3 - it^2\Big|_0^1 = 1 - i \quad \text{and} \quad \int_0^1 e^{2\pi it}\, dt = \left.\frac{e^{2\pi it}}{2\pi i}\right|_0^1 = 0.$$

The second fundamental theorem of calculus asserts that "if $f:\, [a,b] \to \mathbb{R}$ is continuous, then the indefinite integral

$$F(t) = \int_a^t f(x)\, dx, \quad a \le t \le b,$$

is an antiderivative for $f(t)$. Moreover, each antiderivative for $f(t)$ differs from $F(t)$ by a constant." Our next theorem is a two-dimensional analogue of the first fundamental theorem of calculus.

Theorem 7.31. (Green's Theorem) *Let $P(x,y)$ and $Q(x,y)$ be continuous with continuous partials in a simply connected closed region R whose boundary is the contour C. Then*

$$\int_C P\, dx + Q\, dy = \int\!\!\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy, \tag{7.17}$$

where C is traversed in the positive sense.

Proof. We prove the theorem in the special case that R is a rectangle (and its interior) whose sides are parallel to the coordinate axes (see Figure 7.11). Let $C = C_1 + C_2 + C_3 + C_4$ in Figure 7.11. Observing that $dy \equiv 0$ on C_1 and C_3 while $dx \equiv 0$ on C_2 and C_4, we have

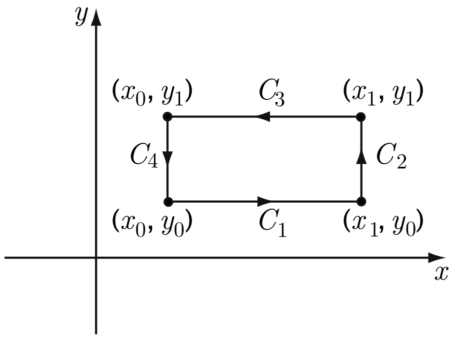


Figure 7.11.

$$\int_C P(x,y)\,dx + Q(x,y)\,dy = \int_{x_0}^{x_1} P(x,y_0)\,dx + \int_{y_0}^{y_1} Q(x_1,y)\,dy + \int_{x_1}^{x_0} P(x,y_1)\,dx + \int_{y_1}^{y_0} Q(x_0,y)\,dy.$$

Combining these integrals, we obtain

$$\int_C P\,dx + Q\,dy = \int_{x_0}^{x_1} \{P(x,y_0) - P(x,y_1)\}\,dx + \int_{y_0}^{y_1} \{Q(x_1,y) - Q(x_0,y)\}\,dy. \tag{7.18}$$

The fundamental theorem of calculus may now be applied to the integrands on the right side of (7.18). This yields

$$\begin{aligned} \int_C P\,dx + Q\,dy &= \int_{x_0}^{x_1}\int_{y_0}^{y_1} -\frac{\partial P}{\partial y}\,dy\,dx + \int_{y_0}^{y_1}\int_{x_0}^{x_1} \frac{\partial Q}{\partial x}\,dx\,dy \quad (7.19)\\ &= \int_{y_0}^{y_1}\int_{x_0}^{x_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy \\ &= \int\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy, \end{aligned}$$

where the interchange in the order of the integration in (7.7) may be understood by viewing the iterated integral as representing a volume. This proves the theorem for the rectangle in Figure 7.11. For a complete proof of Green's theorem, see Apostol [Ap]. ■

Example 7.32. Let us evaluate the line integral

$$\int_C xy\,dx + (x^2 + y^2)\,dy$$

along the square $0 \le x \le 1$, $0 \le y \le 1$. A direct proof gives

$$\begin{aligned}\int_C xy\,dx + (x^2+y^2)\,dy &= \int_0^1 x\cdot 0\,dx + \int_0^1 (1+y^2)\,dy \\ &\quad + \int_1^0 x\cdot 1\,dx + \int_1^0 (0+y^2)\,dy \\ &= \int_0^1 dy - \int_0^1 x\,dx = \frac{1}{2}.\end{aligned}$$

Alternately, by Green's theorem,

$$\int_C xy\,dx + (x^2+y^2)\,dy = \int_0^1\int_0^1 (2x-x)\,dx\,dy = \frac{1}{2}. \qquad \bullet$$

Enough playing around. We now return to complex variables to show the reason for introducing Green's theorem.

Theorem 7.33. (Cauchy's "Weak" Theorem) *If $f(z)$ is analytic (with a continuous derivative) in a simply connected domain D, and C is closed contour lying in D, then we have $\int_C f(z)\,dz = 0$.*

Proof. Set $f(z) = u(x,y) + iv(x,y)$. By the Cauchy–Riemann equations for analytic functions,

$$u_x = v_y, \quad u_y = -v_x \quad \text{for} \quad (x,y) \in D. \tag{7.20}$$

Since $f'(z)$ is presumed continuous, the four partials must also be continuous. Suppose, for the moment, that C is a simple closed contour. Then an application of Green's theorem to (7.8) yields

$$\begin{aligned}\int_C f(z)\,dz &= \int_C u\,dx - v\,dy + i\int_C v\,dx + u\,dy \\ &= \int\int_R \left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dx\,dy + i\int\int_R \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\,dy,\end{aligned}$$

where R is the region enclosed by C. In view of (7.20), both integrands on the right are identically zero in R. This proves the theorem when C is simple closed contour.

For a general closed contour, the proof follows in like manner from a more general statement of Green's theorem. See Apostol [Ap]. ■

Corollary 7.34. *Under the conditions of Theorem 7.33, let C_1 and C_2 be any contours in the domain with the same initial and terminal points. Then*

$$\int_{C_1} f(z)\,dz = \int_{C_2} f(z)\,dz.$$

Proof. Suppose C_1 and C_2 both have initial and terminal points z_0 and z_1 respectively (see Figure 7.12). Let $C = C_1 - C_2$. Then

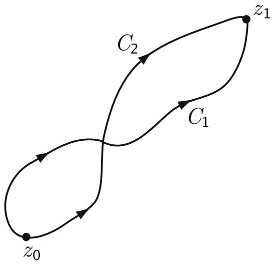


Figure 7.12.

$$\begin{aligned}\int_C f(z)\,dz &= \int_{C_1-C_2} f(z)\,dz \\ &= \int_{C_1} f(z)\,dz + \int_{-C_2} f(z)\,dz \\ &= \int_{C_1} f(z)\,dz - \int_{C_2} f(z)\,dz.\end{aligned}$$

Since C is a closed contour, $\int_C f(z)\,dz = 0$, from which we conclude that $\int_{C_1} f(z)\,dz = \int_{C_2} f(z)\,dz$. ■

Remark 7.35. Corollary 7.34 (as well as Theorem 7.33) says that the integral is independent of path in the domain. That is, the value of the integral just depends on the initial and terminal points, provided only that the contour stays inside the domain where the function is continuously differentiable. Hence, under the conditions of the theorem, we can give meaning to the expression $\int_{z_0}^{z_1} f(z)\,dz$. Its value may be found by computing the complex line integral $\int_C f(z)\,dz$ along any contour C in the domain that has initial point z_0 and terminal point z_1. In particular, if the contour C is closed ($z_1 = z_0$), then

$$\int_C f(z)\,dz = \int_{z_0}^{z_1} f(z)\,dz = 0. \qquad \bullet$$

Cauchy's theorem, in its present form, is weak because the analytic function was required to have a continuous derivative (so that Green's theorem could be applied). While this may seem like a minor restriction, it does not allow us to apply Cauchy's theorem to the class of *all* analytic functions. However, in the next section, this restrictive hypothesis will be eliminated. Then, in Chapter 8, it will be shown that every analytic function does, in fact, have a continuous derivative.

We will now examine the extent to which Cauchy's theorem is valid for multiply connected regions. Recall that

$$\int_{|z|=1} \frac{1}{z}\,dz = \int_0^{2\pi} \frac{ie^{i\theta}}{e^{i\theta}}\,d\theta = 2\pi i.$$

Thus even though the function $f(z) = 1/z$ is analytic everywhere on the unit circle, the above integral is not zero. Note that f has derivatives of all orders in $\mathbb{C} \setminus \{0\}$. Cauchy's theorem is not applicable because the punctured disk is not simply connected. To trace what went wrong with this function, observe that $1/z$ is the derivative of many different branches of $\log z$. Suppose we start with any point w on the unit circle and integrate counterclockwise around the circle through one revolution. While the terminal point $we^{2\pi i}$ has the same location in the plane as the initial point w, its argument has increased by 2π. Choosing a specific branch for the logarithm, with a branch cut on the ray $\arg z = -\alpha$, $w = e^{i\alpha}$, we may now write

$$\int_{|z|=1} \frac{1}{z}\, dz = \int_{|z|=1} \frac{d}{dz} \log z = \log z \Big|_{w}^{we^{2\pi i}} = \log we^{2\pi i} - \log w.$$

This last expression simplifies to

$$\ln |we^{2\pi i}| + i \arg we^{2\pi i} - (\ln |w| + i \arg w) = i(\arg we^{2\pi i} - \arg w) = 2\pi i.$$

Thus the integral is nonzero because the function $1/z$ has many antiderivatives. The value of the integral is related to the change in the argument of the multiple-valued function $\log z$. Note also that the value of the integral is independent of the choice of the initial branch. For simply connected domains, this problem does not arise because analytic functions then have single-valued antiderivatives as we shall see. Moreover, this idea can be extended to any closed contour $C : z(t)$, $a \leq t \leq b$ that does not pass through the origin. Indeed, if C is a closed contour that avoids the origin, then we have

$$\int_C \frac{dz}{z} = \log z(t)\big|_a^b = i \arg z(t)\big|_a^b = i\, \theta|_C$$

where θ is the angle which the line segment $[0, z(t)]$ joining 0 to the variable point $z(t)$ makes with the horizontal line. Thus, the total variation is 2π times the number of times z winds around 0 as z traverses C. In other words,

$$\frac{1}{2\pi i} \int_C \frac{dz}{z}$$

is an integer which is called the *winding number* of C with respect to the origin (see for example [A, P1]).

Suppose we integrate $1/z$ along the boundary C of the multiply connected region consisting of the annulus $r_0 < |z| < r_1$ ($r_0 > 0$). If the integration is performed in the positive sense (where the domain always remains on the left) as shown in Figure 7.13, then

$$\begin{aligned}
\int_C \frac{1}{z}\, dz &= \oint_{|z|=r_1} \frac{1}{z}\, dz + \oint_{|z|=r_0} \frac{1}{z}\, dz \\
&= \int_0^{2\pi} \frac{ir_1 e^{i\theta}}{r_1 e^{i\theta}}\, d\theta + \int_0^{-2\pi} \frac{ir_0 e^{i\theta}}{r_0 e^{i\theta}}\, d\theta \\
&= 2\pi i - 2\pi i = 0.
\end{aligned}$$

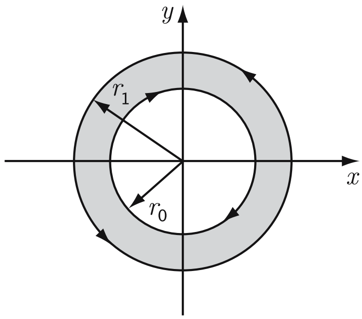


Figure 7.13.

The fact that this integral is zero is a consequence of the following Cauchy theorem for multiply connected regions.

Theorem 7.36. *Suppose that $f(z)$ is analytic (with a continuous derivative) in a multiply connected domain and on its boundary C. Then we have $\int_C f(z)\,dz = 0$, where the integration is performed along C in the positive sense.*

We indicate a method of proof that involves "transforming" a multiply connected region into a simply connected region. To illustrate, consider the multiply connected region in Figure 7.14. Suppose we construct the line segment AB, called a *cross-cut*, which connects the outer boundary C_1 with the inner boundary C_2. Then the domain bounded by the contour C_1, the line segment AB, the contour C_2, and the line segment BA (traversed as illustrated in Figure 7.14) is simply connected. This is so because no closed curve in the new region is allowed to cross the line segment AB. Let C denote the boundary of this domain. Then by Cauchy's theorem for simply connected regions, we have

$$\begin{aligned}\int_C f(z)\,dz &= \oint_{C_1} f(z)\,dz + \int_{AB} f(z)\,dz + \oint_{C_2} f(z)\,dz + \int_{BA} f(z)\,dz \\ &= 0.\end{aligned}$$

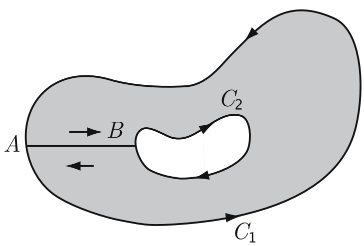


Figure 7.14.

Note that

$$\int_{AB} f(z)\,dz = -\int_{BA} f(z)\,dz,$$

so that

$$\int_{C_1+C_2} f(z)\,dz = \oint_{C_1} f(z)\,dz + \oint_{C_2} f(z)\,dz = 0.$$

This finishes the discussion of Cauchy's theorem for the domain in Figure 7.14. In Figure 7.15, we illustrate Cauchy's theorem for domain with $(n-1)$ holes.

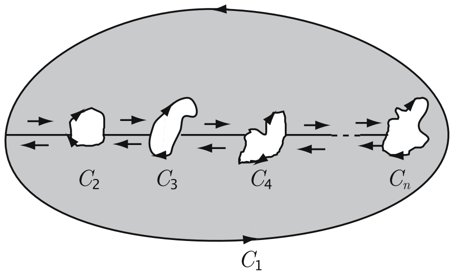


Figure 7.15.

In a manner similar to that used for one hole, we get

$$\int_{C_1+C_2+\cdots+c_n} f(z)\,dz = 0. \tag{7.21}$$

Equation (7.21) can be written in the form

$$\begin{aligned}\oint_{C_1} f(z)\,dz &= -\left[\oint_{C_2} f(z)\,dz + \cdots + \oint_{C_n} f(z)\,dz\right]\\ &= \oint_{-C_2} f(z)\,dz + \cdots + \oint_{-C_n} f(z)\,dz.\end{aligned}$$

In other words, by integrating along each inner contour in the counterclockwise direction, so that the $(n-1)$ inner contours have negative orientation, it follows that the value of the integral along the outer contour is equal to the sum of the values along the inner contours.

In a more complicated multiply connected region, it may not be possible to connect an inner boundary to an outer boundary by a straight line segment; but a polygonal line can always be found that furnishes us with the necessary cross-cut for any multiply connected region. In fact, Green's theorem can also be generalized from simply to multiply connected regions, thus affording us with a direct proof of Cauchy's theorem for multiply connected regions.

Finally, we remark that requiring analyticity on the boundary C in Theorem 7.36 means that the function is actually analytic in a domain containing C.

Questions 7.37.

1. What are the differences between real and complex line integrals? Between a line integral and a Riemann integral?
2. Where was continuity of the partials used in proving Green's theorem for rectangles?
3. Why is Green's theorem a two-dimensional analog to the fundamental theorem of calculus?
4. In Green's theorem, if C were traversed in the negative direction, what could we conclude?
5. How does Green's theorem give us a way to compute the area of a region?
6. Suppose $f(z)$ has a continuous derivative in a simply connected region whose boundary is C. May Cauchy's theorem be applied to conclude $\int_C f(z)\,dz = 0$?
7. Suppose $\int_C f(z)\,dz = 0$ for some contour C. Can anything be said about $f(z)$?
8. Does Cauchy's theorem apply to a function having a continuous derivative in a region exterior to a disk?
9. Can Cauchy's theorem be used to evaluate Riemann integrals?
10. Let $u(z) = u(x, y)$ be a real-valued harmonic function on the unit disk Δ, and γ be a simple closed contour in Δ. Is $\int_\gamma u(z)\,dz = 0$? How about if Δ is replaced by a general domain D?
11. Is $\int_C z\,dz$ independent of the path C between 0 and $1 + i$?
12. Is $\int_C (\operatorname{Re} z)\,dz$ independent of the path C between 0 and $1 + i$?
13. Is $\int_C \overline{z}\,dz$ independent of the path C between 0 and $1 + i$?

Exercises 7.38.

1. Evaluate the following line integrals:
   (a) $\displaystyle\int_C xy\,dx + (x^2+y^2)\,dy$ along the quarter-circle C in the first quadrant having radius $r = 2$.
   (b) $\displaystyle\int_C x^2y\,dx + (2x+1)y^2\,dy$ along the square having vertices $(1, 0)$, $(1, -1)$, $(2, -1)$, and $(2, 0)$.
   (c) $\displaystyle\int_C y^2\,dx + x^2\,dy$ along the curve C parameterized by $x = a\cos^3 t$, $y = a\sin^3 t$ $(0 \le t \le 2\pi)$.
   (d) $\displaystyle\int_C \frac{xy^2}{x^2+y^2}\,dy$ along the circle $|z| = r$.
   (e) $\displaystyle\int_C (x^2 + xy)\,dy$ along the parabola $y = x^2$ from $(-2, 4)$ to $(2, 4)$.
2. Let C be any simple closed contour bounding a region having area A. Prove that

$$A = \frac{1}{2}\int_C x\,dy - y\,dx = -\int_C y\,dz = -i\int_C x\,dz = -\frac{i}{2}\int_C \overline{z}\,dz.$$

3. Modify the proof of Green's theorem for a rectangle to show that

$$\int_C P(x,y)\,dx = -\int\int_R \frac{\partial P}{\partial y}\,dx\,dy, \quad \int_C Q(x,y)\,dy = \int\int_R \frac{\partial Q}{\partial x}\,dx\,dy.$$

4. Verify Cauchy's theorem for the functions $3z-2$ and z^2+3z-1 if C is the square having corners $\pm 1 \pm i$.
5. By evaluating $\int_{|z|=1} e^z\,dz$, show that

$$\int_{-\pi}^{\pi} e^{\cos\theta}\cos(\theta+\sin\theta)\,d\theta = \int_{-\pi}^{\pi} e^{\cos\theta}\sin(\theta+\sin\theta)\,d\theta = 0.$$

6. Evaluate the following integrals along any contour between the points represented by the limits of integration.

(a) $\displaystyle\int_{-\pi i}^{\pi i} e^z\,dz$ (b) $\displaystyle\int_0^{\pi+i} e^{iz}\,dz$ (c) $\displaystyle\int_{1-i}^{2+i} (z^2+3z-2)\,dz.$

7.4 Cauchy's Theorem

The central theme of this section is to investigate conditions to cover general situations so that the integral of an analytic function along a closed contour vanishes. We will actually prove several forms of Cauchy's theorem (also called the Cauchy–Goursat theorem), each involving different geometric and topological considerations. Goursat showed that Theorem 7.33 can be proved without assuming the continuity of $f'(z)$. In its simplest form, the theorem is proved for a rectangle. The proof involves a construction similar to that used in the proof of the Heine–Borel theorem (Theorem 2.26). The ultimate aim is to understand precisely the local structure of analytic functions.

Theorem 7.39. (Cauchy's theorem for a rectangle) *Let $f(z)$ be analytic in a domain containing a rectangle C and its interior. Then $\int_C f(z)\,dz = 0$.*

Proof. Divide C into four congruent rectangles $C^{(1)}$, $C^{(2)}$, $C^{(3)}$ and $C^{(4)}$ as indicated in Figure 7.16, and let $I_1^j = \int_{C^{(j)}} f(z)\,dz$ for $1 \le j \le 4$. The integrals over the common sides have opposite orientation, and hence cancel one another. Therefore, from the known properties in the complex integral, it follows that

$$I := \int_C f(z)\,dz = \sum_{j=1}^{4} I_1^j.$$

By the triangle inequality,

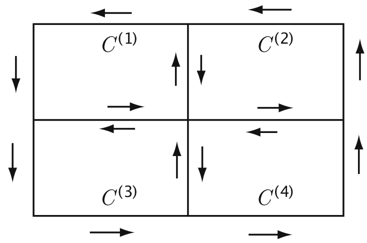


Figure 7.16.

$$|I| \le \sum_{j=1}^{4} |I_1^j|. \tag{7.22}$$

If every term in this sum were less than $|I|/4$, then we would get a contradiction. Thus, for at least one of the terms on the right side of (7.22), denoted conveniently by $I_1 = \int_{C_1} f(z)\,dz$, we have

$$|I_1| \ge |I|/4.$$

Next divide the rectangle C_1 into four congruent rectangles, and, as above, observe that for at least one, denoted by C_2

$$|I_2| = \left|\int_{C_2} f(z)\,dz\right| \ge \frac{|I_1|}{4} \ge \frac{|I|}{4^2}.$$

Continuing the process, we obtain a nested sequence of rectangles $\{C_n\}$ (see Figure 7.17), each satisfying the inequality

$$|I_n| = \left|\int_{C_n} f(z)\,dz\right| \ge \frac{|I_{n-1}|}{4} \ge \cdots \ge \frac{|I|}{4^n}$$

so that

$$|I| \le 4^n |I_n| \quad \text{for } n = 1, 2, 3, \ldots . \tag{7.23}$$

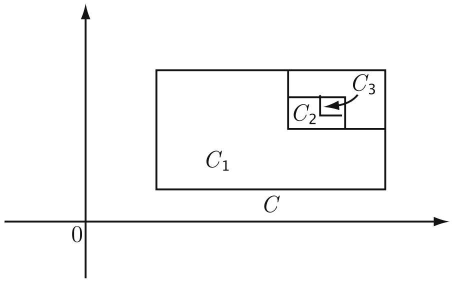


Figure 7.17.

According to Lemma 2.25, there is exactly one point, call it z_0, belonging to all the rectangles. For each n, the point z_0 is either on or inside the rectangle C_n.

In particular, z_0 is in the domain of analyticity of $f(z)$. Hence given $\epsilon > 0$, there exists a $\delta > 0$ such that

$$f'(z_0) = \frac{f(z) - f(z_0)}{z - z_0} - \eta(z), \tag{7.24}$$

where $|\eta(z)| < \epsilon$ when $|z - z_0| < \delta$. Solving (7.24) for $f(z)$ and integrating, we get

$$\begin{aligned} \int_{C_n} f(z)\,dz &= \int_{C_n} \{f(z_0) + f'(z_0)(z - z_0) + \eta(z)(z - z_0)\}\,dz \qquad (7.25) \\ &= \int_{C_n} \{f(z_0) + f'(z_0)(z - z_0)\}\,dz \\ &\qquad + \int_{C_n} \eta(z)(z - z_0)\,dz, \end{aligned}$$

which is valid for each n. The integrand of the first integral has a continuous derivative in the entire complex plane. Thus Cauchy's weak theorem may be applied to obtain

$$\int_{C_n} \{f(z_0) + f'(z_0)(z - z_0)\}\,dz = 0.$$

Therefore, (7.25) simplifies to

$$I_n = \int_{C_n} f(z)\,dz = \int_{C_n} \eta(z)(z - z_0)\,dz.$$

Now choose n large enough so that $C_n \subset N(z_0; \delta)$. Then

$$\begin{aligned} |I_n| &= \left| \int_{C_n} \eta(z)(z - z_0)\,dz \right| \\ &\leq \int_{C_n} |\eta(z)|\,|z - z_0|\,|dz| \leq \epsilon \int_{C_n} |z - z_0|\,|dz|. \end{aligned}$$

Denote the length of the diagonal and the perimeter of C_n by D_n and L_n, respectively. Then $|z - z_0| \leq D_n$ for all z in C_n, and

$$|I_n| \leq \epsilon D_n L_n = \epsilon \frac{D}{2^n}\frac{L}{2^n} = \frac{\epsilon DL}{4^n}, \tag{7.26}$$

where D and L denote, respectively, the length of the diagonal and the perimeter of C. Combining (7.26) with (7.23), we obtain

$$|I| \leq 4^n \frac{\epsilon DL}{4^n} = \epsilon DL.$$

Since ϵ was arbitrary, $|I| = |\int_C f(z)\,dz| = 0$ and the proof is complete. ■

For instance, by Theorem 7.39, we have

$$\int_C \frac{2z^2+1}{(z+3)^5(z^2+8)z^3(z-1)}\,dz = 0$$

where C is the positively oriented square with vertices $1+i$, $1+3i$, $2+3i$, $2+i$.

Corollary 7.40. *Let f be continuous in a domain D containing a rectangle C and its interior. Suppose that f is analytic in $D\setminus\{a\}$ for some point $a \in D$. Then $\int_C f(z)\,dz = 0$.*

Proof. It suffices to prove the theorem when a lies inside C. As before divide C into n^2 congruent rectangles C_{jk} (see Figure 7.18 for illustration when $n = 4$). From the elementary properties of complex line integrals, we have

$$\int_C f(z)\,dz = \sum_{j=1}^{n}\sum_{k=1}^{n}\int_{C_{jk}} f(z)\,dz.$$

If a is neither an interior point nor a point of C_{jk}, then, by Theorem 7.39, $\int_{C_{jk}} f(z)\,dz = 0$. On the other hand, if a is inside or on the rectangle C_{jk}, then the M-L inequality shows that

$$\left|\int_{C_{jk}} f(z)\,dz\right| \le \int_{C_{jk}} |f(z)|\,|dz| \le M\,L(C_{jk}) = \frac{ML(C)}{n},$$

where $M = \max_{z\in C}|f(z)|$, $L(C_{jk})$ and $L(C)$ represent the perimeter of C_{jk} and C, respectively. Note that $|f(z)|$ is a continuous function on the compact set C, and the point a at the worst can belong to one of the four rectangles C_{jk}. It follows that

$$\left|\int_C f(z)\,dz\right| = \left|\sum_{a\in C_{jk}}\int_{C_{jk}} f(z)\,dz\right| \le \sum_{a\in C_{jk}}\left|\int_{C_{jk}} f(z)\,dz\right| \le \frac{4ML(C)}{n}.$$

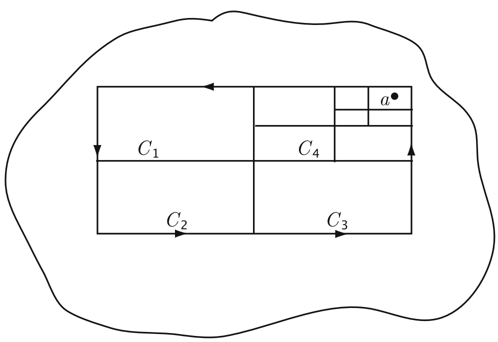


Figure 7.18.

Since n was arbitrary, $\left|\int_C f(z)\,dz\right| = 0$ and the proof is complete. ∎

Let us pause for a moment to summarize what we have shown and where we are headed. In the previous section, it was shown that for a function having a continuous derivative in a domain, the integral around any closed contour in the domain is zero, or, equivalently, the integral along any contour in the domain depends only on the end points of the contour. The previous theorem eliminates the requirement of continuity for the derivative when the contour is a rectangle.

Our goal is to show that the rectangle in Theorem 7.39 may be replaced by an arbitrary closed contour in the domain. This will be accomplished by first showing that every continuous function having an antiderivative in a domain also has the property that the integral is independent of path. Next we will show that a function analytic in a disk has an antiderivative, and then that a function analytic in a simply connected domain has an antiderivative. Finally, Cauchy's theorem will be extended to multiply connected domains by "transforming" them into simply connected domains, as was done in the previous section. We start with the following theorem which is an analogue of the first fundamental theorem of calculus.

Theorem 7.41. (Fundamental Theorem of Integration) *Let $f(z)$ be continuous in a domain D, and suppose there is a differentiable function $F(z)$ such that $F'(z) = f(z)$ in D. Then for any contour C in D parameterized by $z(t)$, $a \le t \le b$, we have*

$$\int_C f(z)\,dz = F(z(b)) - F(z(a)).$$

In particular, if C is closed then $\int_C f(z)\,dz = 0$.

Proof. Since $F(z)$ has a continuous derivative in D, we get

$$\begin{aligned}
\int_C f(z)\,dz &= \int_C F'(z)\,dz \\
&= \int_a^b F'(z(t))z'(t)\,dt \\
&= \int_a^b \frac{d}{dt}(F(z(t))\,dt = F(z(b)) - F(z(a)),
\end{aligned}$$

the last equality following from the fundamental theorem of integral calculus. If we use a more familiar notation $z(a) = z_0$ and $z(b) = z_1$, then the conclusion may be expressed as

$$\int_{z_0}^{z_1} f(z)\,dz = F(z_1) - F(z_0)$$

along any contour C in the domain having initial point z_0 and terminal point z_1. If the contour is closed, then $z(a) = z(b) = z_0$ so that

$$\int_C f(z)\,dz = \int_{z_0}^{z_0} f(z)\,dz = F(z_0) - F(z_0) = 0. \qquad \blacksquare$$

Moreover, Theorem 7.41 is a consequence of the corresponding formula for line integrals. Indeed, as $F'(z) = F_x = -iF_y$, it follows that

$$\begin{aligned} F(z_1) - F(z_0) &= \int_{z_0}^{z_1} dF \\ &= \int_{z_0}^{z_1} F_x\,dx + F_y\,dy \\ &= \int_{z_0}^{z_1} F'(z)(dx + i\,dy) \\ &= \int_{z_0}^{z_1} F'(z)\,dz = \int_{z_0}^{z_1} f(z)\,dz. \end{aligned}$$

Example 7.42. The function $f(z) = z^n$ $(n \in \mathbb{N})$ is continuous everywhere and has an antiderivative $F(z) = z^{n+1}/(n+1)$. Hence for any contour C in the plane from z_0 to z_1,

$$\int_C z^n\,dz = \int_{z_0}^{z_1} z^n\,dz = \frac{z_1^{n+1}}{n+1} - \frac{z_0^{n+1}}{n+1}.$$

In particular, if C is a closed curve $(z_0 = z_1)$ then for each $n \in \mathbb{N}$, we have that $\int_C f(z)\,dz = 0$. More generally, if $p(z) = \sum_{k=0}^n a_k z^k$ is a polynomial, then

$$P(z) = \sum_{k=0}^{n} \frac{a_k}{k+1} z^{k+1} + c$$

is primitive of $p(z)$, $P'(z) = p(z)$, and so

$$\int_{z_0}^{z_1} p(z)\,dz = P(z_1) - P(z_0).$$

In particular, $\int_C p(z)\,dz = 0$ if C is closed curve in $\mathbb{C}$. ●

Examples 7.43. By Theorem 7.41, we obtain the following:

(i) Clearly, $\int_{|z|=1} \csc^2 z\,dz = 0$. Indeed if $f(z) = \csc^2 z$, then $F(z) = -\cot z$ has the property that $F'(z) = \csc^2 z$ and $F(z)$ is analytic in $\mathbb{C} \setminus \{n\pi : n \in \mathbb{Z}\}$. In particular, f and F are analytic for $0 < |z| < \pi$. Similarly, we obtain that
$$\int_{|z|=1} \sec^2 z\,dz = 0.$$

(ii) Suppose we wish to evaluate $\int_C (z+a)e^{bz}\,dz$ $(b \neq 0)$, where C is the parabolic arc $x^2 = y$ from $(0,0)$ to $(1,1)$. First we note that if $f(z) = (z+a)e^{bz}$, then $F(z)$ for which $F'(z) = f(z)$ is given by

$$F(z) = (z+a)\frac{e^{bz}}{b} - \frac{e^{bz}}{b^2} = \frac{e^{bz}}{b^2}(b(z+a)-1)$$

(which may be obtained by integrating $(z+a)e^{bz}$ by parts). Thus, by Theorem 7.41, we have

$$\int_C (z+a)e^{bz}\,dz = F(1+i) - F(0).$$

Similarly,

$$\int_0^1 (z+a)e^{bz}\,dz = F(1) - F(0) = (1+a)\frac{e^b}{b} - \frac{e^b}{b^2} - \frac{a}{b} + \frac{1}{b^2}.$$

(iii) If C is the quarter circle $|z| = 2$ in the first quadrant joining 2 to $2i$, then, according to Theorem 7.41, we have

$$\int_C z^n\,dz = \left.\frac{z^{n+1}}{n+1}\right|_2^{2i} = \frac{2^{n+1}}{n+1}(i^{n+1}-1) \quad (n \in \mathbb{Z}\setminus\{-1\}).$$ •

Theorem 7.41 looks deceptively similar to the fundamental theorem of integral calculus. There is an important difference. The fundamental theorem says that a continuous function $f(x)$ defined on $[a,b]$ *has* an antiderivative $F(x)$ satisfying

$$\int_{x_0}^{x_1} f(t)\,dt = F(x_1) - F(x_0) \quad (a \le x_0 < x_1 \le b).$$

Theorem 7.41 merely asserts that if the continuous function $f(z)$ has an antiderivative, then the conclusion follows. That continuity is not a sufficient condition for the existence of an antiderivative can be seen by the following example.

Example 7.44. If the everywhere continuous function $f(z) = \overline{z}$ had an antiderivative, then the conclusion of Theorem 7.41 would follow. But

$$\int_{|z|=1} \overline{z}\,dz = \int_{-\pi}^{\pi} e^{-it} i e^{it}\,dt = 2\pi i \neq 0.$$

This shows that $f(z) = \overline{z}$ does not have an antiderivative. •

However, as seen in Example 7.42, Theorem 7.41 provides a powerful tool for evaluating definite integrals. So, in order to evaluate $\int_{z_0}^{z_1} f(z)\,dz$, it suffices to find a analytic function $F(z)$ such that $F'(z) = f(z)$. But finding such an $F(z)$ is not always easy. For instance, if $f(z) = \sin(1/z)$ or $\cos(1/z)$, how do we know whether $F(z)$ exists or what precisely is $F(z)$?

We now examine the relationship between antiderivatives and analytic functions. The following theorem, at least on the local level provides a condition which guarantees existence of the antiderivatives of a function (see also Theorem 7.39).

Theorem 7.45. (Cauchy's theorem for a disk) *Let $f(z)$ be analytic in a domain containing the closed disk $|z-z_0| \leq r$. Then $\int_{|z-z_0|=r} f(z)\,dz = 0$.*

Proof. In view of Theorem 7.41, it suffices to find a function $F(z)$ such that $F'(z) = f(z)$ for $|z-z_0| \leq r$. Choose any point $z = x+iy$ in the disk and let C_1 be the contour consisting of the horizontal line segment from $z_0 = x_0+iy_0$ to $x+iy_0$ followed by the vertical line segment from $x+iy_0$ to $x+iy$. Also, let C_2 be the contour consisting of the vertical line segment from $z_0 = x_0+iy_0$ to x_0+iy followed by the horizontal line segment from x_0+iy to $x+iy$ (see Figure 7.19).

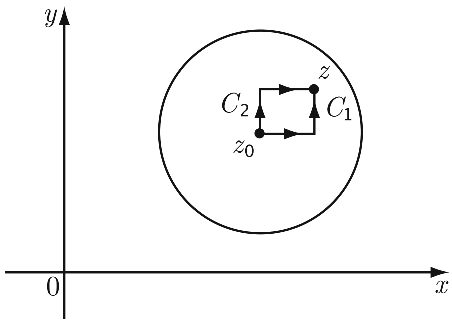


Figure 7.19.

By Theorem 7.39 and basic properties of integrals,

$$\int_{C_1-C_2} f(z)\,dz = \int_{C_1} f(z)\,dz - \int_{C_2} f(z)\,dz = 0. \tag{7.27}$$

Define

$$F(z) = \int_{C_1} f(z)\,dz = \int_{x_0}^{x} f(t+iy_0)\,dt + \int_{y_0}^{y} f(x+it)i\,dt. \tag{7.28}$$

In view of (7.27), $F(z)$ may also be expressed as

$$F(z) = \int_{C_2} f(z)\,dz = \int_{y_0}^{y} f(x_0+it)i\,dt + \int_{x_0}^{x} f(t+iy)\,dt. \tag{7.29}$$

Taking the partial derivative of $F(z)$ with respect to y in (7.28), we obtain (since the first term in right side of (7.28) is independent of y)

$$\frac{\partial F}{\partial y} = i\frac{\partial}{\partial y}\left(\int_{y_0}^{y} f(x+it)\,dt\right) = if(x+iy) = if(z); \tag{7.30}$$

(here the fundamental theorem of calculus is applied to

$$\int_{y_0}^{y} g(x,t)\,dt = \int_{y_0}^{y} u(x,t)\,dt + i\int_{y_0}^{y} v(x,t)\,dt, \quad g = u + iv).$$

Similarly, taking the partial derivative of $F(z)$ with respect to x in (7.29), we have (since the first term in (7.29) is independent of x)

$$\frac{\partial F}{\partial x} = \frac{\partial}{\partial x}\left(\int_{x_0}^{x} f(t+iy)\,dt\right) = f(x+iy) = f(z); \tag{7.31}$$

(again this is a consequence of the fundamental theorem of calculus). In view of (7.30) and (7.31),

$$F_x(z) = -iF_y(z) = f(z). \tag{7.32}$$

But (7.32) is just the Cauchy–Riemann equations for $F(z)$. Furthermore, the continuity of partials F_x and F_y on the disk follows from the continuity of $f(z)$. Hence Theorem 5.17 may be applied to establish the analyticity of F at z. Since z was arbitrary, $F(z)$ is analytic in the disk $|z - z_0| \leq r$. Finally, from (5.5), we conclude that $F'(z) = F_x(z) = f(z)$, i.e., F is a primitive of f in the disk $|z - z_0| \leq r$. ∎

Corollary 7.46. *Let f be analytic for $|z - z_0| < r$ except at some point a inside the disk and continuous for $|z - z_0| \leq r$. Then $\int_{|z-z_0|=r} f(z)\,dz = 0$.*

Proof. Follows if we combine Corollary 7.40 and Theorem 7.45. ∎

A circle has the property that any point inside can be joined to the center by two distinct broken line segments, which, when taken together, form the perimeter of a rectangle whose sides are parallel to the coordinate axes. Furthermore, this is the *only* property that was used in going from Cauchy's theorem for a rectangle to Cauchy's theorem for a circle. In a more general domain, no such construction is possible; however according to Remark 2.1, every pair of points in a domain D can be joined by a polygonal line lying in D (with sides parallel to the coordinate axes). In Ahlfors [A], it is shown that if the domain is simply connected, then two such polygonal lines C_1 and C_2 can be constructed so that their difference $C_1 - C_2$ consists of a finite number of boundaries of rectangles traversed alternately in the positive and negative directions, as illustrated in Figure 7.20.

This fact will be used in proving our main theorem.

Theorem 7.47. (Cauchy's Theorem) *If $f(z)$ is analytic in a simply connected domain D and C is a closed contour lying in D, then $\int_C f(z)\,dz = 0$.*

Proof. According to Theorem 7.41, it suffices to find a function $F(z)$ such that $F'(z) = f(z)$ in the simply connected domain D. Fix a point z_0 in D and choose an arbitrary z in D. Then with C_1 and C_2 constructed as in Figure 7.20, we define $F(z)$ by

$$F(z) = \int_{C_1} f(z)\, dz.$$

According to Theorem 7.31,

$$\int_{C_1 - C_2} f(z)\, dz = \int_{C_1} f(z)\, dz - \int_{C_2} f(z)\, dz = 0,$$

because the integral around each rectangle is zero. Hence, we also have

$$F(z) = \int_{C_2} f(z)\, dz.$$

Suppose $z_1 = x_1 + iy_1$ is the last point of intersection of C_1 and C_2 between z_0 and $z = x + iy$. Also suppose that, in this last rectangle, C_1 consists of the horizontal followed by the vertical line, whereas C_2 consists of the vertical followed by the horizontal, as shown in Figure 7.20.

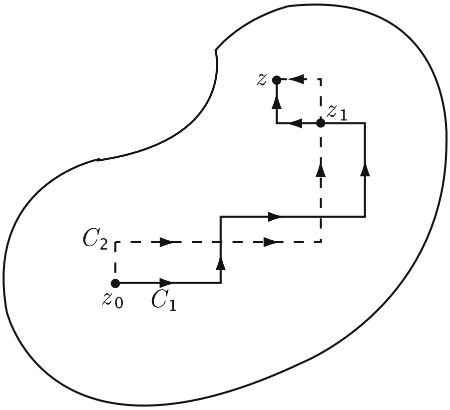


Figure 7.20.

In view of Theorem 7.39, the value for the integral of $f(z)$ from z_0 to z_1 is the same along both the contours C_1 and C_2, and we denote their common value by K. The remainder of the proof is similar to that of Theorem 7.45, for we have

$$F(z) = \int_{C_1} f(z)\, dz = K + \int_{x_1}^{x} f(t + iy_1)\, dt + \int_{y_1}^{y} f(x + it) i\, dt \quad (7.33)$$

and

$$F(z) = \int_{C_2} f(z)\, dz = K + \int_{y_1}^{y} f(x_1 + it) i\, dt + \int_{x_1}^{x} f(t + iy)\, dt. \quad (7.34)$$

From (7.33), we get

$$\frac{\partial F}{\partial y} = \frac{\partial}{\partial y}\left(\int_{y_1}^{y} f(x_1 + it)\,dt\right) = if(x + iy) = if(z)$$

and, from (7.34),

$$\frac{\partial F}{\partial x} = \frac{\partial}{\partial x}\left(\int_{x_1}^{x} f(t + iy)\,dt\right) = f(x + iy) = f(z).$$

Hence

$$F_x(z) = -iF_y(z) = f(z). \tag{7.35}$$

Equation (7.35) represents the Cauchy–Riemann equations for $F(z)$. The partials of $F(z)$ are continuous because $f(z)$ is continuous. Therefore, $F(z)$ is analytic in D, with $F'(z) = F_x(z) = f(z)$ in D. ■

Remark 7.48. The topological notions utilized in the proof of Theorem 7.47 allowed us to deal with finitely many rectangles inside the simply connected domain, from whence Theorem 7.39 was applicable. But the essence of Theorem 7.39 consisted of "shrinking" a rectangle to a point. Consequently, Cauchy's theorem ultimately relies on the fact that $\int_{z_0}^{z_0} f(z)\,dz = 0$. ●

Theorem 7.36, which generalized Cauchy's "weak" theorem (Theorem 7.33) from simply to multiply connected domains, was purely topological in nature, and nowhere used the continuity of the partials. Hence, Cauchy's theorem is also valid for a multiply connected region, the proof consisting of "transforming" a multiply connected region into a simply connected region, as in the proof of Theorem 7.36. We remark that if the contour encloses singularities of the function, we cannot use Cauchy's theorem. For example, consider

$$I = \int_C \frac{1}{(z-1)^2}\,dz$$

along a simple closed contour having the point 1 as an interior point. Note that $F(z) = -1/(1-z)$ is an antiderivative of $f(z) = (z-1)^{-2}$ for $z \in \mathbb{C}\setminus\{1\}$. According to Theorem 7.41, $I = 0$. But Theorem 7.47 is not applicable. On the other hand, for example, if

$$f(z) = \frac{\sin z}{(z-1)^2} \quad \text{or} \quad \frac{e^z}{(z-1)^2},$$

then it is not clear whether these functions have antiderivatives. To cover a situation like this, we need to develop another theorem called Cauchy's integral formula which we shall do in Chapter 8. Its extension in the form of the Residue theorem will be discussed in Chapter 9. On the other hand, for certain situations the following theorem is helpful to simplify the problem by replacing the given contour by another, or others.

Theorem 7.49. (Cauchy's Theorem for multiply connected domains) *Let D be a multiply connected domain bounded externally by a simple closed contour C and internally by n simple closed nonintersecting contours C_1, $C_2, \ldots, C_n$. Let f be analytic on $D \cup C \cup C_1 \cup C_2 \cup \cdots \cup C_n$. Then*

$$\int_C f(z)\,dz = \sum_{k=1}^{n} \int_{C_k} f(z)\,dz$$

where C is taken counterclockwise around the external boundary C and clockwise around the internal boundaries $C_1, C_2, \ldots, C_n$.

This generalization of Cauchy's theorem aids us in evaluating integrals along a contour enclosing a region in which the function is not analytic.

First we evaluate $\int_C (z-z_0)^{-1}\,dz$ along a simple closed contour C having z_0 is an interior point.

For some $\epsilon > 0$, the circle $C : |z - z_0| = \epsilon$ is interior to the contour C. Also, the function $f(z) = 1/(z - z_0)$ is analytic in the multiply connected region between C and $|z - z_0| = \epsilon$ (see Figure 7.21). Hence

$$0 = \int_{C+C_1} \frac{1}{z-z_0}\,dz = \oint_C \frac{1}{z-z_0}\,dz + \oint_{C_1} \frac{1}{z-z_0}\,dz.$$

Therefore,

$$\oint_C \frac{1}{z-z_0}\,dz = -\oint_{C_1} \frac{1}{z-z_0}\,dz = \oint_{-C_1} \frac{1}{z-z_0}\,dz.$$

Note that the positive orientation of C_1 is clockwise so that the positive orientation of $-C_1$ is counterclockwise. Parameterizing $-C_1$ by $z(t) = \epsilon e^{it}$, $0 \le t \le 2\pi$, we have

$$\oint_C \frac{1}{z-z_0}\,dz = \oint_{-C_1} \frac{1}{z-z_0}\,dz = \int_0^{2\pi} \frac{z'(t)}{z(t)}\,dt = \int_0^{2\pi} \frac{i\epsilon e^{it}}{\epsilon e^{it}}\,dt = 2\pi i.$$

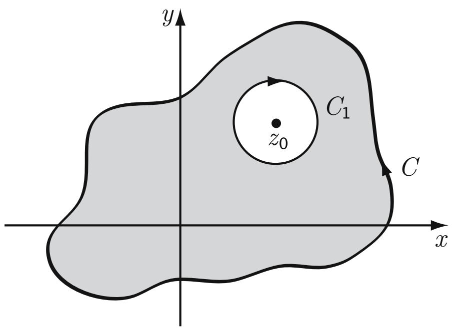


Figure 7.21.

Thus the validity of Cauchy's theorem for multiple connected regions takes the worry out of parameterizing ugly contours. The method used in the above example shows that $\oint_C (z-z_0)^{-1}\,dz = 2\pi i$ or 0 according to whether the point z_0 is inside or outside the simple closed contour C. No additional information is required in order to evaluate the integral. On the other hand, for $n \in \mathbb{N}\backslash\{1\}$ the function $f(z) = (z-z_0)^{-n}$ is not analytic at z_0 and for the same contour C, we can easily see that

$$\oint_C \frac{1}{(z-z_0)^n}\,dz = 0.$$

Observe that this result does not contradict Theorem 7.47. Note also that in all these cases, the value of the integral does not depend on the radius ϵ, as long as C_1 lies inside C.

We now illustrate Theorem 7.49 by evaluating the integral

$$I = \int_{|z|=4} f(z)\,dz, \quad f(z) = \frac{1}{z^2+4}.$$

Letting $C_1 = \{z : |z-2i| = 1\}$ and $C_2 = \{z : |z+2i| = 1\}$, we have, by Theorem 7.49,

$$\begin{aligned} I &= \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz \\ &= \frac{1}{4i}\int_{C_1}\left(\frac{1}{z-2i} - \frac{1}{z+2i}\right)dz + \frac{1}{4i}\int_{C_2}\left(\frac{1}{z-2i} - \frac{1}{z+2i}\right)dz \\ &= \frac{1}{4i}(2\pi i + 0) + \frac{1}{4i}(0 - 2\pi i) = 0. \end{aligned}$$

Similarly, one can easily show that

$$\int_C \frac{dz}{z^2+r^2} = 0 = \int_C \frac{dz}{z^2-r^2} \quad (r>0)$$

where C is the positively oriented circle $|z| = r + 1/2$.

Example 7.50. We wish to evaluate the integral

$$\int_{|z|=2} \frac{dz}{z^3-3}.$$

This integral may be evaluated without using Cauchy's residue theorem which will be discussed in Chapter 9. Define $D = \{z : 2 < |z| < R\}$. Then, $f(z) = 1/(z^3-3)$ is analytic in D for each $R > 2$. By Cauchy's theorem for multiply connected domains,

$$\int_{|z|=2} \frac{dz}{z^3-3} = \int_{|z|=R} \frac{dz}{z^3-3}$$

and the value of the integral must be independent of R. By the M-L Inequality,

$$\left| \int_{|z|=R} \frac{dz}{z^3 - 3} \right| \leq \frac{2\pi R}{R^3 - 3} \to 0 \quad \text{as } R \to \infty$$

which shows that the value of the given integral is 0. The same argument may be used for similar integrals. For example, we see that

$$\int_{|z|=2} \frac{dz}{1 + z + z^2 + z^3} = 0.$$

Note that $(1 + z + z^2 + z^3)^{-1} = (1 - z)/(1 - z^4)$. ∎

Cauchy's theorem for multiply connected regions cannot be proved directly by the same method as was used for simply connected regions, because analytic functions need not have (single-valued) analytic antiderivatives in multiply connected domains. In the above example, $\log(z - z_0)$ is the analytic antiderivative of $1/(z - z_0)$ only when confined to a branch. This concept of *analytic logarithm* in simply connected domains is made more explicit in the following theorem.

Theorem 7.51. *If $f(z)$ is analytic and nonzero in a simply connected domain D, then there exists a function $g(z)$, analytic in D, such that $e^{g(z)} = f(z)$.*

Proof. Since $f(z)$ never vanishes in D, the function $f'(z)/f(z)$ is analytic in D. Furthermore, the integral of $f'(z)/f(z)$ between any two points in D is independent of the path in the simply connected domain D. We define $g(z)$ by the formula

$$g(z) = \int_{z_0}^{z} \frac{f'(\zeta)}{f(\zeta)}\, d\zeta + \operatorname{Log} f(z_0), \tag{7.36}$$

where z_0 is fixed point in D, z is an arbitrary point in D, and the path of integration is any path that lies in D. Set $h(z) = f(z)e^{-g(z)}$, and observe that

$$\begin{aligned} h'(z) &= f'(z)e^{-g(z)} - f(z)g'(z)e^{-g(z)} \\ &= f'(z)e^{-g(z)} - f(z)\frac{f'(z)}{f(z)}e^{-g(z)} \\ &= 0. \end{aligned}$$

Thus $h(z)$ is a constant in D. To determine the constant, we set $z = z_0$ to obtain

$$h(z_0) = f(z_0)e^{-g(z_0)} = f(z_0)e^{-\operatorname{Log} f(z_0)} = 1.$$

Therefore, $f(z)e^{-g(z)} \equiv 1$ throughout D, and the theorem is proved. ∎

To see that the hypothesis that the domain be simply connected is essential, observe that $1/z$ never vanishes in the punctured plane and cannot be expressed as $e^{g(z)}$ for an *analytic* function $g(z)$. (Recall that no branch of $-\log z$ is analytic in the punctured plane.)

Corollary 7.52. *If $f(z)$ is analytic and nonzero in a simply connected domain D, then an analytic branch of $(f(z))^{1/n}$ (n a positive integer) can be defined in D.*

Proof. Set $f(z) = e^{g(z)}$, where $g(z)$ is analytic in D, and the existence of g provided in Theorem 7.51. Define the nth root function $(f(z))^{1/n}$ by $(f(z))^{1/n} = e^{(1/n)g(z)}$. ■

Remark 7.53. More generally, each of the n functions $e^{(g(z)+2k\pi i)/n}$ ($k = 0, 1, 2, \ldots, n-1$) is an analytic branch of $(f(z))^{1/n}$. ●

We end this section with an example.

Example 7.54. Consider

$$\int_{|z|=1} f(z)\,dz, \quad f(z) = 1/z^{1/2}.$$

Then $z = 0$ is a branch point of $f(z)$. If we choose principal branch, then

$$z^{1/2} = e^{(1/2)\,\mathrm{Log}\,z} = e^{(1/2)(\ln|z|+i\mathrm{Arg}\,z)}$$

so that for $z = e^{i\theta}$, we have $z^{1/2} = e^{i(1/2)\mathrm{Arg}\,z} = e^{(1/2)i\theta}$ and

$$\int_{-\pi}^{\pi} \frac{ie^{i\theta}}{e^{i\theta/2}}\,d\theta = i\int_{-\pi}^{\pi} e^{-i\theta/2}\,d\theta = 4i.$$

If C is the line segment $[1, 1+i]$ connecting 1 and $1+i$, and if we choose the principal branch for $z^{1/2}$, then we have $F(z) = z^{1/2} = e^{(1/2)\,\mathrm{Log}\,z}$, and so

$$F'(z) = e^{(1/2)\,\mathrm{Log}\,z}\frac{1}{2z} = \frac{1}{2}\frac{e^{(1/2)\,\mathrm{Log}\,z}}{e^{\mathrm{Log}\,z}} = \frac{1}{2}\frac{1}{e^{(1/2)\,\mathrm{Log}\,z}} = \frac{1}{2}f(z)$$

where F is analytic on $\mathbb{C}\backslash(-\infty, 0]$ with $F'(z) = f(z)$. Using this we compute that

$$\int_C \frac{dz}{z^{1/2}} = 2\int_C F'(z)\,dz = 2\,[F(1+i) - F(1)] = 2\,[2^{1/4}e^{i\pi/8} - 1]. \qquad ●$$

Questions 7.55.

1. If $f(z)$ is analytic in a domain and C is a closed contour in the domain, does $\int_C |f(z)|\,dz = 0$? Does $\int_C f(z)\,|dz| = 0$?
2. Where in the proof of Theorem 7.47 did we use the fact that the domain was simply connected?
3. Suppose $f(z)$ is analytic on a contour C. Does $\int_C f(z)\,dz = 0$?
4. If f is continuous on the contour C, is $\int_C f(z)\,dz = -\int_{-C} f(z)\,dz$?

5. For what type of simple closed contours, is

$$\int_\gamma (1 + z + z^2 + \cdots + z^n)^{-1}\, dz = 0?$$

6. What is the relationship between Theorem 7.41 and Green's theorem?
7. Where, in the proof of Theorem 7.45, was the hypothesis of analyticity needed?
8. What are the differences between Cauchy's theorem and Cauchy's weak theorem?
9. What can be said about $\int_C (1/z)\, dz$ if the contour C passes through the origin?
10. What values may be assumed by $\int_C (1/z)\, dz$ if C is a closed curve that is not simple?
11. Is the function $g(z)$ in Theorem 7.51 unique?
12. What is an antiderivative of $\cos(z^2)$? of $\sin(z^2)$? Is it possible to use Theorem 7.41 to conclude that $\int_C \sin(z^2)\, dz = 0$ for any simple closed contour C?
13. What is a domain D of analyticity of $f(z) = z(z-1)^{1/2}$? Find an antiderivative of f in D?
14. When can a contour integral $\int_C f(z)\, dz$ be independent of the path?
15. What is a complex version of the fundamental theorem of calculus?
16. Let C be a closed contour. Does $\int_C z\, dz = 0$? Does $\mathrm{Re} \int_C \overline{z}\, dz = 0$? Does $\int_C (\mathrm{Re}\, z)\, dz = 0$? Does $\mathrm{Im} \int_C \overline{z}\, dz = 0$? Does $\int_C (\mathrm{Im}\, z)\, dz = 0$?

Exercises 7.56.

1. Evaluate $\int_{-i}^{i} |z|\, dz$ along different contours. Does $|z|$ have an antiderivative?
2. Evaluate $\int_\gamma f(z)\, dz$, where
   (i) $f(z) = z^3$ and $\gamma(t) = t^2 + it$ for $t \in [0, \pi]$
   (ii) $f(z) = \sin z$ and $\gamma(t) = t + it^2$ for $t \in [0, \pi/2]$
   (iii) $f(z) = 1/z$ and $\gamma(t) = \cos t^2 - i \sin t$ for $t \in [0, \pi/2]$
   (iv) $f(z) = 1/z$ and $\gamma(t) = -\cos t - ie \sin t$ for $t \in [0, \pi/2]$.
3. Give an example of a function $f(z)$ for which $\int_{|z|=r} f(z)\, dz = 0$ for each $r > 0$ even though $f(z)$ is not analytic everywhere.
4. Let $f = u + iv$ be analytic inside and on a simple closed contour γ. Show by an example that Cauchy's theorem does not hold separately for the real and imaginary parts of f.
5. Find $\int_C (1 + z^2)^{-1}\, dz$, where C is the circle

   (a) $|z - i| = 1$ (b) $|z + i| = 1$ (c) $|z| = 2$ (d) $|z - 1| = 1$.

6. Separate the integrand into real and imaginary parts and evaluate, where possible, the expression $\int_C (e^z/z)\, dz$, where C is

   (a) $|z| = 1$ (b) $|z-2| = 1$ (c) the square having vertices $\pm 1 \pm i$.

7. Suppose $\operatorname{Re} z_0 > 0$ and $\operatorname{Re} z_1 > 0$. Evaluate $\int_{z_0}^{z_1} (1/z)\, dz$ along contours in the right half-plane.
8. Let $a, b \in \mathbb{C}$ and $r > 0$. Then, by decomposing the integrand into partial fractions, show that

$$\int_{|z|=r} \frac{dz}{(z-a)(z-b)} = \begin{cases} 0 & \text{if } |a| > r \text{ and } |b| > r \ (a, b \neq 0) \\ 0 & \text{if } |a| < r \text{ and } |b| < r \ (a, b \in \mathbb{C}) \\ \dfrac{2\pi i}{a-b} & \text{if } |a| < r < |b| \\ \dfrac{2\pi i}{b-a} & \text{if } |b| < r < |a|. \end{cases}$$

9. Suppose that f is analytic for $|z| < 2$ and α is a complex constant. Evaluate
$$I = \int_{|z|=1} (\operatorname{Re} z + \alpha) \frac{f(z)}{z}\, dz.$$
10. Evaluate $\int_{z_1}^{z_2} a^z\, dz$ ($a \neq 0$ is given).
11. Find $\int_{|z|=1} f(z)\, dz$, when

    (a) $f(z) = (z \sin z)/(z+2) + \overline{z}$ (b) $f(z) = z^4 + iz + 2\operatorname{Im} z$.

8

Applications of Cauchy's Theorem

Most of the powerful and beautiful theorems proved in this chapter have no analog in real variables. While Cauchy's theorem is indeed elegant, its importance lies in applications. In this chapter, we prove several theorems that were alluded to in previous chapters. We prove the Cauchy integral formula which gives the value of an analytic function in a disk in terms of the values on the boundary. Also, we show that an analytic function has derivatives of all orders and may be represented by a power series. The fundamental theorem of algebra is proved in several different ways. In fact, there is such a nice relationship between the different theorems in this chapter that it seems any theorem worth proving is worth proving twice.

8.1 Cauchy's Integral Formula

If $f(z)$ is analytic in a simply connected domain D, then we know already that $\int_C f(z)\,dz = 0$ along every closed contour C contained in D. An interesting variation occurs when a function is analytic at all but a finite number of points. As we have seen in the previous chapter,

$$\int_C \frac{1}{z - z_0}\,dz = 2\pi i \tag{8.1}$$

along every positively oriented simple closed contour C containing z_0. We now develop the Cauchy integral formula which is indeed a generalization of (8.1). Moreover, Cauchy's integral formula leads to three important properties of analytic functions that are unparalleled in real variable methods:

- every analytic function is infinitely differentiable, see Theorem 8.3;
- every analytic function can be expressed locally as a Taylor series in the vicinity of a point of analyticity, see Theorem 8.8;
- every analytic function can be expressed as a Laurent series in the vicinity of an isolated singularity, see Section 9.2.

These facts hinge on the following result.

Theorem 8.1. (Cauchy's First Integral Formula) *Let $f(z)$ be analytic in a simply connected domain containing the simple closed contour C. If z_0 is inside C, then*

$$f(z_0) = \frac{1}{2\pi i}\int_C \frac{f(z)}{z-z_0}\,dz.$$

Proof. Given $\epsilon > 0$, construct a circle $C_r : |z - z_0| = r$ inside C and small enough so that $|f(z) - f(z_0)| < \epsilon$ for all z on C_r. According to Cauchy's theorem for multiply connected regions,

$$\int_C \frac{f(z)}{z-z_0}\,dz = \int_{C_r} \frac{f(z)}{z-z_0}\,dz.$$

Thus, writing $f(z) = f(z_0) + (f(z) - f(z_0))$ for the integral on the right, one has

$$\begin{aligned}\int_C \frac{f(z)}{z-z_0}\,dz &= \int_{C_r} \frac{f(z_0)}{z-z_0}\,dz + \int_{C_r} \frac{f(z)-f(z_0)}{z-z_0}\,dz \\ &= 2\pi i f(z_0) + \int_{C_r} \frac{f(z)-f(z_0)}{z-z_0}\,dz, \quad \text{since } \int_{C_r} \tfrac{1}{z-z_0}\,dz = 2\pi i.\end{aligned}$$

Since the integral in the left side has a fixed value, as does $2\pi i f(z_0)$, it follows that for each C_r, the value of the contour integral

$$\int_{C_r} \frac{f(z)-f(z_0)}{z-z_0}\,dz$$

is constant. We now show that this value must be zero. We have

$$\left|\int_{C_r} \frac{f(z)-f(z_0)}{z-z_0}\,dz\right| \le \int_{C_r} \frac{|f(z)-f(z_0)|}{|z-z_0|}\,|dz| < \frac{\epsilon}{r}(2\pi r) = 2\pi\epsilon.$$

Since ϵ is arbitrary, the integral is zero. This concludes the proof. ∎

Remark 8.2. When $f(z) \equiv 1$, the conclusion of the theorem reduces to (8.1). Moreover, Theorem 8.1 expresses the value of $f(z)$ at any point *inside* C in terms of its values *on* C. In other words, if $f(z)$ is known to be analytic inside and on the boundary of a simply connected domain, then the values of $f(z)$ on the boundary completely determine the values of $f(z)$ inside. There is no analog to this theorem for functions of a real variable. More precisely, when a real-valued function $f(x)$ is differentiable on the closed interval $[a, b]$, its value at $x = a$ and $x = b$ in no way can dictate the value of $f(x)$ on the open interval (a, b). For instance, for each $n \in \mathbb{N}$, the functions

$$f_n(x) = x^n, \;\; 0 \le x \le 1,$$

all have the same boundary values $(f(0) = 0, \; f(1) = 1)$ but differ from one another at all interior points. Similarly, we see that when a function $f(x, y)$ of two real variables $x, \; y$ real differentiable inside and on a simple closed contour C, its value on C do not determine the values of $f(x, y)$ inside C. ●

We next express the derivative of an analytic function in terms of an integral. Using the notation of Theorem 8.1, choose h small enough in absolute value so that $z_0 + h$ is inside C. Then,

$$f(z_0) = \frac{1}{2\pi i}\int_C \frac{f(z)}{z - z_0}\,dz \quad \text{and} \quad f(z_0 + h) = \frac{1}{2\pi i}\int_C \frac{f(z)}{z - (z_0 + h)}\,dz.$$

Hence, for $h \neq 0$,

$$\begin{aligned} \frac{f(z_0 + h) - f(z_0)}{h} &= \frac{1}{2\pi i h}\int_C \left(\frac{1}{z - z_0 - h} - \frac{1}{z - z_0}\right) f(z)\,dz \qquad (8.2) \\ &= \frac{1}{2\pi i}\int_C \frac{f(z)}{(z - z_0 - h)(z - z_0)}\,dz. \end{aligned}$$

When $h \to 0$, the integrand approaches $f(z)/(z - z_0)^2$. It appears probable that the limit is

$$\frac{1}{2\pi i}\int_C \frac{f(z)}{(z - z_0)^2}\,dz,$$

although in general the limit of the integrand is not necessarily the same as the integrand of the limit. To prove that we may take the limit inside the integral, we must show that the difference

$$\begin{aligned} &\left|\frac{1}{2\pi i}\int_C \frac{f(z)}{(z - z_0 - h)(z - z_0)}\,dz - \frac{1}{2\pi i}\int_C \frac{f(z)}{(z - z_0)^2}\,dz\right| \qquad (8.3) \\ &\qquad = \left|\frac{h}{2\pi i}\int_C \frac{f(z)}{(z - z_0 - h)(z - z_0)^2}\,dz\right| \end{aligned}$$

can be made arbitrarily small. Given a circle $C_1 : |z - z_0| = r$ contained in C, choose h small enough so that $|h| \leq r/2$ (see Figure 8.1). Note that $z_0 \in \operatorname{Int}(C_1)$ and

$$|z_0 + h - z_0| < r/2 \quad \Longleftrightarrow |h| < r/2,$$

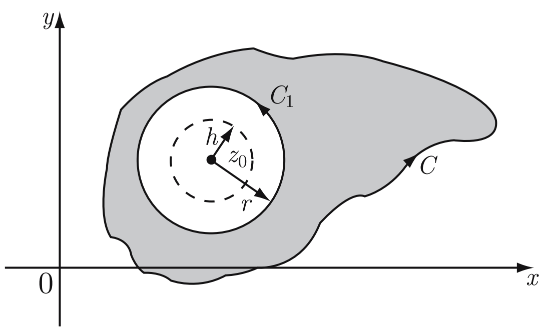


Figure 8.1.

which shows that $z_0 + h \in \operatorname{Int}(C_1)$. Further, for $z \in C_1$, we have

$$|z - z_0 - h| \geq |z - z_0| - |h| \geq r - |h| \geq r - r/2 > 0$$

and so, by Cauchy's theorem for multiply connected domains, we have

$$\int_C \frac{f(z)}{(z - z_0 - h)(z - z_0)^2}\, dz = \int_{C_1} \frac{f(z)}{(z - z_0 - h)(z - z_0)^2}\, dz.$$

Since $f(z)$ is continuous on C_1, it is bounded (say $|f(z)| \leq M$ on C_1). Thus

$$\begin{aligned}
\left| \frac{h}{2\pi i} \int_{C_1} \frac{f(z)}{(z - z_0 - h)(z - z_0)^2}\, dz \right| &\leq \frac{|h|}{2\pi} \int_{C_1} \frac{|f(z)|}{|z - z_0 - h|\, |z - z_0|^2}\, |dz| \\
&\leq \frac{|h|M}{2\pi r^2} \int_{C_1} \frac{|dz|}{|z - z_0| - |h|} \\
&\leq \frac{|h|M}{\pi r^3} \int_{C_1} |dz| \\
&= \frac{|h|M}{\pi r^3} (2\pi r) = |h| \left(\frac{2M}{r^2} \right).
\end{aligned}$$

Clearly, the limit of this expression as $h \to 0$ is zero and so, (8.3) tends to zero with h. In view of (8.2) and (8.3),

$$\begin{aligned}
f'(z_0) &= \lim_{h \to 0} \frac{1}{2\pi i} \int_C \frac{f(z)}{(z - z_0 - h)(z - z_0)}\, dz \\
&= \frac{1}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^2}\, dz.
\end{aligned} \tag{8.4}$$

Equation (8.4) expresses the value of $f'(z)$ at any point inside C in terms of the values of $f(z)$ on C. Moreover, (8.4) shows that the operations of differentiation and contour integration can be interchanged. We knew, by hypothesis, that $f(z)$ was differentiable at all points inside C. But the above process can be repeated. From (8.4), we obtain

$$\frac{f'(z_0 + h) - f'(z_0)}{h} = \frac{1}{2\pi i} \int_C \frac{\{2(z - z_0) - h\} f(z)}{(z - z_0)^2 (z - z_0 - h)^2}\, dz. \tag{8.5}$$

As before, we can show that the limit may be taken inside the integral so that (8.5) leads to

$$f''(z_0) = \frac{2}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^3}\, dz. \tag{8.6}$$

Equation (8.6) gives much more information than we had a right to expect. First, we have the *existence* of $f''(z)$ at all points inside C. Next, the second derivative may be expressed in terms of the values of $f(z)$ on C. This argument can be repeated indefinitely. An induction shows that the nth derivative is given by

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz. \tag{8.7}$$

To see this, we assume that (8.7) holds for $n = k \geq 1$. Then

$$f^{(k)}(z_0) = \frac{k!}{2\pi i} \int_C \frac{f(z)}{(z-z_0)^{k+1}}\, dz.$$

We must show that $f^{(k+1)}(z_0)$ exists and the formula (8.7) holds for $n = k+1$. To do this, we use the binomial expansion

$$(z-z_0-h)^{k+1} = (z-z_0)^{k+1} - (k+1)(z-z_0)^k h + \frac{(k+1)k}{2}(z-z_0)^{k-1}h^2 - \cdots$$

to express

$$\begin{aligned}
&\frac{1}{(z-(z_0+h))^{k+1}} - \frac{1}{(z-z_0)^{k+1}} \\
&\quad = \frac{(z-z_0)^{k+1} - (z-z_0-h)^{k+1}}{(z-z_0)^{k+1}(z-z_0-h)^{k+1}} \\
&\quad = \frac{(k+1)h}{(z-z_0)(z-z_0-h)^{k+1}} - \frac{(k+1)k}{2} \frac{h^2}{(z-z_0)^2(z-z_0-h)^{k+1}} + \cdots
\end{aligned}$$

where the dots indicate terms with powers of h up to h^{k+1}. In view of this expression,

$$\frac{f^{(k)}(z_0+h) - f^{(k)}(z_0)}{h} = \frac{k!}{2\pi i} \int_C \frac{f(z)}{z-z_0} \left[\frac{k+1}{(z-z_0-h)^{k+1}} + O(h) \right] dz.$$

Letting $h \to 0$, we obtain (8.7) for $n = k+1$,

$$f^{(k+1)}(z_0) = \frac{(k+1)!}{2\pi i} \int_C \frac{f(z)}{(z-z_0)^{k+2}}\, dz.$$

Hence, every analytic function has derivatives of all orders and derivatives of all orders at each point may be expressed in terms of the values of the function on its boundary. We sum up this remarkable result with

Theorem 8.3. (Generalized Cauchy's Integral Formula) *Let $f(z)$ be analytic in a simply connected domain containing the simple closed contour C. Then $f(z)$ has derivatives of all orders at each point z_0 inside C, with*

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz.$$

Remark 8.4. While Theorem 8.3 is stated as a global property, a property of simply connected domains, it is really a local property. Since a function analytic at a point is also analytic in a (simply connected) neighborhood of the point, it follows that a function analytic at a point must have derivatives of all orders at that point. ●

Note how this result is radically different from the theory of functions of a real variable. In basic calculus, we have learned that the existence of the derivative of $f(x)$ does not guarantee the continuity of the derivative $f'(x)$, much less the differentiability of $f'(x)$, see the example on p. 134. Now, we consider the function

$$f(x) = x^{7/5}$$

which has a first derivative for all $x \in \mathbb{R}$ and

$$f'(x) = (7/5)x^{2/5}.$$

Observe that $f'(x)$ does not have a first derivative at $x = 0$ and therefore, $f(x)$ does not have a second derivative at the origin. Similarly, $f(x) = x^{11/5}$ has a first and second derivative on $\mathbb{R}$ but has no third derivative at $x = 0$. It follows that the existence of n derivatives of a real-valued function $f(x)$ does not guarantee the $(n+1)$th derivative of $f(x)$. Thus, Theorem 8.3 does not hold in the case of functions of a real variable.

This example demonstrated one essential difference between functions of a complex variable and functions of a real variable.

Remark 8.5. Cauchy's integral formula is also valid for multiply connected regions. We prove it for the multiply connected region in Figure 8.2. In the exercises, the reader is asked to supply the general proof. ●

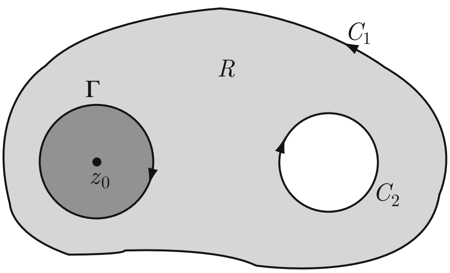


Figure 8.2.

Suppose $f(z)$ is analytic in the multiply connected region R whose boundary consists of the contour $C = C_1 \cup C_2$. Construct a circle Γ contained in R and having center at z_0. Then by Cauchy's theorem for multiply connected regions,

$$\frac{1}{2\pi i}\oint_{C_1}\frac{f(z)}{z-z_0}\,dz+\frac{1}{2\pi i}\oint_{C_2}\frac{f(z)}{z-z_0}\,dz+\frac{1}{2\pi i}\oint_{\Gamma}\frac{f(z)}{z-z_0}\,dz=0.$$

Thus, in view of Theorem 8.1,

$$\begin{aligned}f(z_0)&=\frac{1}{2\pi i}\oint_{\Gamma}\frac{f(z)}{z-z_0}\,dz\\&=\frac{1}{2\pi i}\oint_{C_1}\frac{f(z)}{z-z_0}\,dz-\frac{1}{2\pi i}\oint_{C_2}\frac{f(z)}{z-z_0}\,dz.\end{aligned}$$

That is,

$$f(z_0)=\frac{1}{2\pi i}\int_C\frac{f(z)}{z-z_0}\,dz,$$

and this proves Theorem 8.1 for the multiply connected region R. Similarly, we can show that the conclusion of Theorem 8.3 remains valid for the multiply connected region R (see Figure 8.2). Next we have an immediate and important corollary to Theorem 8.3.

Corollary 8.6. *If $f = u + iv$ is analytic in a domain D, then all partial derivatives of u and v exist and are continuous in D.*

Proof. Let $f = u + iv$ be analytic in D. Then, by (5.3),

$$f'(z)=u_x+iv_x=v_y-iu_y.$$

By the analyticity of $f'(z)$, it follows that each of the first partial derivatives of u and v exist and are continuous in D because $f'(z)$ is continuous in D. Because $f'(z)$ is analytic in D, from the above equation, we again have

$$\begin{aligned}f''(z)&=u_{xx}+iv_{xx}\\&=(v_x)_y-i(u_x)_y\\&=(v_y)_x-i(u_y)_x.\end{aligned}$$

This process may be continued to conclude that u and v have continuous partial derivatives of all orders at each point where the function $f = u + iv$ is analytic. ■

Example 8.7. Setting $f(z) = z - 3\cos z$, we compute

$$\int_{|z|=2}\frac{z-3\cos z}{(z-\pi/2)^2}\,dz=2\pi i f'\left(\frac{\pi}{2}\right)=8\pi i. \qquad \bullet$$

We are now able to prove Taylor's theorem for complex functions.

Theorem 8.8. (Taylor's Theorem) *Let $f(z)$ be analytic in a domain D whose boundary is C. If z_0 is a point in D, then $f(z)$ may be expressed as*

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$$

and the series converges for $|z - z_0| < \delta$*, where* δ *is the distance from* z_0 *to the nearest point on* C*.*

Proof. Construct a circle C_1 having center at z_0 and radius ρ, $\rho < \delta$. Let z be any point inside C_1. By Cauchy's integral formula,

$$f(z) = \frac{1}{2\pi i} \int_{C_1} \frac{f(\zeta)}{\zeta - z}\, d\zeta. \tag{8.8}$$

Set $|z - z_0| = r$. Then $r = |z - z_0| < |\zeta - z_0| = \rho$ (see Figure 8.3), and

$$\begin{aligned} \frac{1}{\zeta - z} &= \frac{1}{\zeta - z_0}\left(\frac{1}{1 - (z - z_0)/(\zeta - z_0)}\right) \\ &= \frac{1}{\zeta - z_0}\left(1 + \frac{z - z_0}{\zeta - z_0} + \left(\frac{z - z_0}{\zeta - z_0}\right)^2 + \cdots \right. \\ &\qquad \left. + \left(\frac{z - z_0}{\zeta - z_0}\right)^{n-1} + \frac{((z - z_0)/(\zeta - z_0))^n}{1 - (z - z_0)/(\zeta - z_0)}\right). \end{aligned} \tag{8.9}$$

In view of (8.8), we may multiply (8.9) by $f(\zeta)$ and integrate to obtain

$$\begin{aligned} f(z) = \frac{1}{2\pi i} \int_{C_1} \frac{f(\zeta)}{\zeta - z_0}\, d\zeta + \frac{(z - z_0)}{2\pi i} \int_{C_1} \frac{f(\zeta)}{(\zeta - z_0)^2}\, d\zeta + \\ + \cdots + \frac{(z - z_0)^{n-1}}{2\pi i} \int_{C_1} \frac{f(\zeta)}{(\zeta - z_0)^n}\, d\zeta + R_n, \end{aligned} \tag{8.10}$$

where

$$R_n = \frac{1}{2\pi i} \int_{C_1} \left(\frac{z - z_0}{\zeta - z_0}\right)^n \frac{f(\zeta)}{\zeta - z}\, d\zeta.$$

But by Cauchy's integral formula, (8.10) may be expressed in the form

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \cdots + \frac{f^{(n-1)}(z_0)}{(n-1)!}(z - z_0)^{n-1} + R_n.$$

The result follows if we can show that the remainder term R_n approaches zero as n approaches ∞. Suppose $|f(z)| \le M$ on C_1 (see Figure 8.3). Then

$$R_n \le \frac{1}{2\pi} \int_{C_1} \left|\frac{z - z_0}{\zeta - z_0}\right|^n \left|\frac{f(\zeta)}{\zeta - z}\right| |d\zeta| \le \frac{M}{2\pi}\left(\frac{r}{\rho}\right)^n \int_{C_1} \frac{1}{|\zeta - z|}|d\zeta|. \tag{8.11}$$

Starting with the inequality

$$\frac{1}{|\zeta - z|} = \frac{1}{|\zeta - z_0 - (z - z_0)|} \le \frac{1}{|\zeta - z_0| - |z - z_0|} = \frac{1}{\rho - r},$$

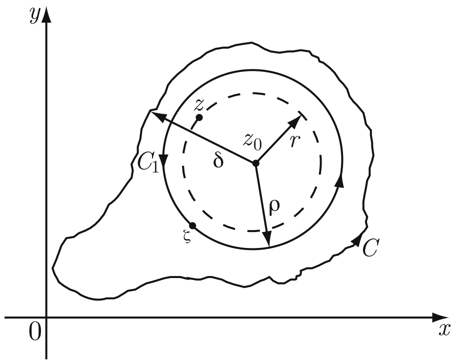


Figure 8.3.

(8.11) leads to

$$R_n \leq \frac{M}{2\pi(\rho - r)} \left(\frac{r}{\rho}\right)^n \int_{C_1} |d\zeta| = \frac{M\rho}{\rho - r} \left(\frac{r}{\rho}\right)^n.$$

Since $r < \rho$, $(r/\rho)^n \to 0$ as $n \to \infty$. Thus,

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n,$$

and the proof is complete. ∎

Remark 8.9. By the M-test (Theorem 6.31), the Taylor series for $f(z)$ also converges uniformly on compact subsets of $|z - z_0| < \delta$. ●

In Section 6.3, we saw that a power series represents an analytic function inside its circle of convergence. Theorem 8.8 is essentially the converse. Thus, a function $f(z)$ is analytic at a point z_0 if and only if $f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n$ in some disk $|z - z_0| \leq r$, where

$$a_n = \frac{f^{(n)}(z_0)}{n!} = \frac{1}{2\pi i} \int_{|z-z_0|=r} \frac{f(z)}{(z - z_0)^{n+1}}\, dz.$$

Theorem 8.8 justifies the Maclaurin expansion

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad \sin z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n-1)!} z^{2n-1}, \quad \cos z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}$$

that were stated, without proof, in Chapter 6 (see also the Examples below).

Examples 8.10. (i) Consider $f(z) = \sin z$. Then f is entire. Also, for each $n \in \mathbb{N}$, we have $f'(z) = \cos z = \sin(z + \pi/2)$ and

$$f''(z) = \cos(z + \pi/2) = \sin(z + \pi/2 + \pi/2) = \sin(z + \pi).$$

Consequently, $f^{(n)}(z) = \sin(z + n\pi/2)$ and so

$$f^{(n)}(0) = \sin(n\pi/2) = \begin{cases} 0 & \text{if } n = 2k \\ (-1)^k & \text{if } n = 2k+1 \end{cases}, \quad k \in \mathbb{N}_0.$$

It follows that

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)!} z^{2n-1}, \quad z \in \mathbb{C}. \tag{8.12}$$

A similar method gives that

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots = \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}, \quad z \in \mathbb{C}. \tag{8.13}$$

Also, it is much easier to use (8.12), and the relation $\frac{d}{dz}(\sin z) = \cos z$ to achieve (8.13), because of Theorem 6.51.

(ii) Suppose we wish to find the Taylor expansion of $f(z) = e^z$ about the point $z = 1$. We could of course, begin by computing the coefficients using the formula for a_n in Theorem 8.8. To avoid this, we may simply rewrite

$$f(z) = e^{z-1+1} = ee^{z-1} = e\sum_{n=0}^{\infty} \frac{(z-1)^n}{n!}, \quad |z-1| < \infty,$$

because of the known series expansion for e^z.

(iii) To find the Taylor expansion for $f(z) = ze^z$ about a point $z = a$, we may simply rewrite

$$\begin{aligned} f(z) &= (z-a+a)e^{z-a+a} \\ &= e^a[(z-a)e^{z-a} + ae^{z-a}] \\ &= e^a\left[\sum_{n=0}^{\infty} \frac{(z-a)^{n+1}}{n!} + a + a\sum_{n=1}^{\infty} \frac{(z-a)^n}{n!}\right], \quad |z-a| < \infty, \\ &= e^a\left[\sum_{n=1}^{\infty} \frac{(z-a)^n}{(n-1)!} + a + a\sum_{n=1}^{\infty} \frac{(z-a)^n}{n!}\right] \\ &= e^a\left[a + \sum_{n=1}^{\infty} \frac{n+a}{n!}(z-a)^n\right], \quad \text{for all } z \in \mathbb{C}. \end{aligned}$$

For instance if $a = 1$, it follows that

$$ze^z = e\left[1 + \sum_{n=1}^{\infty} \frac{n+1}{n!}(z-1)^n\right] \quad \text{for all } z \in \mathbb{C}.$$

(iv) To find the Taylor expansion for $1/z$ about $a \neq 0$, we simply follow the above path, namely,

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

which is valid for $|z-a| < |a|$. For example, for $a = i, -i$, one has

- $\displaystyle \frac{1}{z} = \frac{1}{i}\sum_{n=0}^{\infty}\frac{(-1)^n}{i^n}(z-i)^n := \sum_{n=0}^{\infty} i^{n-1}(z-i)^n, \quad |z-i| < 1,$
- $\displaystyle \frac{1}{z} = \frac{1}{-i}\sum_{n=0}^{\infty}\frac{(-1)^n}{(-i)^n}(z+i)^n = -\sum_{n=0}^{\infty}\frac{(z+i)^n}{i^{n+1}}, \quad |z+i| < 1.$

(v) A similar technique may be adopted to find the Taylor expansion about $z = a \neq 0$, for $f(z) = 1/z^2$. To do this, we first recall that for $|z| < 1$

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n$$

and, because of Theorem 6.51, differentiating with respect to z gives

$$\frac{1}{(1-z)^2} = \sum_{n=1}^{\infty} nz^{n-1} = \sum_{n=0}^{\infty}(n+1)z^n, \quad |z| < 1.$$

Using this, we may write

$$\begin{aligned}\frac{1}{z^2} &= \frac{1}{(a+z-a)^2} = \frac{1}{a^2}\frac{1}{[1+(z-a)/a]^2}\\ &= \frac{1}{a^2}\sum_{n=0}^{\infty}(n+1)(-1)^n\left(\frac{z-a}{a}\right)^n, \quad |z-a| < |a|.\end{aligned}$$

In particular, the Taylor expansion of $1/z^2$ about $a = 1$ follows:

$$\frac{1}{z^2} = \sum_{n=0}^{\infty}(-1)^n(n+1)(z-1)^n, \quad |z-1| < 1.$$

Moreover, it is clear that for $|z-1| < 1$,

$$\frac{z-1}{z^2} = \sum_{n=1}^{\infty}(-1)^{n-1}n(z-1)^n$$

which is the Taylor series for $(z-1)/z^2$ about $z = 1$. We can often build on results such as this. •

We now examine some relationships between uniform convergence and integration.

Theorem 8.11. *Let $\{f_n(z)\}$ be a sequence of functions continuous on a contour C, and suppose that $\{f_n(z)\}$ converges uniformly to $f(z)$ on C. Then*

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

Proof. The statement of the theorem requires us to show that the sequence $\int_C f_n(z)\,dz$ converges to $\int_C f(z)\,dz$.

Note that, by Theorem 6.26, $f(z)$ is continuous on C so that $\int_C f(z)\,dz$ exists. Given $\epsilon > 0$, there is an integer $N = N(\epsilon)$ such that

$$|f_n(z) - f(z)| < \epsilon \ \text{ for } n > N \text{ and all } z \text{ on } C.$$

Denoting the length of C by L, it follows, for $n > N$, that

$$\begin{aligned}\left|\int_C f_n(z)\,dz - \int_C f(z)\,dz\right| &= \left|\int_C (f_n(z) - f(z))\,dz\right| \\ &\le \int_C |f_n(z) - f(z)|\,|dz| < \epsilon L.\end{aligned}$$

Since ϵ is arbitrary, the proof is complete. ■

Corollary 8.12. *Suppose $\{f_n(z)\}$ is a sequence of continuous functions and that $\sum_{n=0}^{\infty} f_n(z)$ converges uniformly on a contour C. Then*

$$\sum_{n=0}^{\infty}\left(\int_C f_n(z)\,dz\right) = \int_C\left(\sum_{n=0}^{\infty} f_n(z)\right)dz.$$

Proof. Set $S_n(z) = \sum_{k=0}^{n} f_k(z)$. Then,

$$\sum_{n=0}^{\infty}\left(\int_C f_n(z)\,dz\right) = \lim_{n\to\infty}\sum_{k=0}^{n}\int_C f_k(z)\,dz = \lim_{n\to\infty}\int_C S_n(z)\,dz.$$

But by Theorem 8.11,

$$\lim_{n\to\infty}\int_C S_n(z)\,dz = \int_C \lim_{n\to\infty} S_n(z) = \int_C\left(\sum_{n=0}^{\infty} f_n(z)\right)dz. \qquad ■$$

Remark 8.13. Our proof of Theorem 8.8 mimicked the proof in the real case. In view of this corollary, we can now give a simpler proof that does not involve a remainder term. Instead of (8.9), we can write

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

the convergence being uniform on C_1. Hence we may multiply by $f(\zeta)/2\pi i$ and integrate term-by-term. This leads directly to

$$
\begin{aligned}
f(z) &= \frac{1}{2\pi i}\int_{C_1} \frac{f(\zeta)}{\zeta - z}\, d\zeta \\
&= \frac{1}{2\pi i}\int_{C_1} \left(\sum_{n=0}^{\infty} \frac{(z-z_0)^n}{(\zeta - z_0)^{n+1}} \right) f(\zeta)\, d\zeta \\
&= \sum_{n=0}^{\infty} (z-z_0)^n \left(\frac{1}{2\pi i}\int_{C_1} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}}\, d\zeta \right) \\
&= \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n.
\end{aligned}
$$

●

We might expect a function $f(z)$ to be analytic at a point z_0 if

$$\int_C f(z)\, dz = 0$$

along every simple closed contour in which z_0 is an interior point. Unfortunately, this converse of Cauchy's theorem is not true. For example,

$$\int_C \frac{1}{z^2}\, dz = 0$$

along every simple closed contour C having the origin as an interior point. This is because $f(z)$ is analytic in the region between C and some circle $|z| = \epsilon$ contained in C. Thus by Cauchy's theorem for multiply connected regions,

$$\int_C \frac{1}{z^2}\, dz = \int_{|z|=\epsilon} \frac{1}{z^2}\, dz = \int_0^{2\pi} \frac{i\epsilon e^{i\theta}}{\epsilon^2 e^{i\theta}}\, d\theta = \frac{i}{\epsilon}\int_0^{2\pi} e^{-i\theta}\, d\theta = 0.$$

But we do have a partial converse to Cauchy's theorem even when the domain D is not simply connected.

Theorem 8.14. (Morera's Theorem) *Let $f(z)$ be continuous in a domain D. If $\int_C f(z)\, dz = 0$ along every simple closed contour C contained in D, then $f(z)$ is analytic in D.*

Proof. Fixing z_0 in D, the value of the function

$$F(z) = \int_{z_0}^{z} f(\zeta)\, d\zeta$$

is independent of the path of integration from z_0 to z inside D. Choose h small enough so that the line segment from z to $(z+h)$ lies in D, and consider the difference quotient

$$\frac{F(z+h) - F(z)}{h} = \frac{1}{h}\left[\int_{z_0}^{z+h} f(\zeta)\, d\zeta - \int_{z_0}^{z} f(\zeta)\, d\zeta \right] = \frac{1}{h}\int_{z}^{z+h} f(\zeta)\, d\zeta.$$

Then

$$\frac{F(z+h)-F(z)}{h} - f(z) = \frac{1}{h}\int_z^{z+h} (f(\zeta) - f(z))\, d\zeta, \tag{8.14}$$

where the path from z to $(z+h)$ is taken to be the straight line segment. By the continuity of $f(z)$, we have $|f(\zeta) - f(z)| < \epsilon$ for $|h|$ sufficiently small. Since the line segment from z to $(z+h)$ has length $|h|$, it follows from (8.14) that

$$\left|\frac{F(z+h)-F(z)}{h} - f(z)\right| < \frac{1}{|h|}\epsilon|h| = \epsilon.$$

Therefore, $F(z)$ is analytic in D with $F'(z) = f(z)$. Moreover, $F(z)$ must have derivatives of all orders. In particular, $F''(z) = f'(z)$ at all points in D, thus proving the analyticity of $f(z)$. ∎

Corollary 8.15. *Let $f(z)$ be analytic in a simply connected domain D. If z_0 is a point in D, then $F(z) = \int_{z_0}^{z} f(\zeta)\, d\zeta$ is analytic in D.*

Proof. According to Cauchy's theorem for simply connected domains, we have $\int_C f(z)\, dz = 0$ along every closed contour C contained in D. Hence $f(z)$ satisfies the hypotheses (as well as the conclusion) of Morera's theorem. The result is thus implicit in the proof of Morera's theorem. ∎

Recall that even if $f(z)$ is analytic in a domain D, we are not guaranteed that $\int_C f(z)\, dz = 0$ along every simple closed contour C contained in D. The function $f(z) = 1/z$ is analytic in the annulus bounded by the circles $|z| = 1/2$ and $|z| = 2$. But $\int_{|z|=1}(1/z)\, dz = 2\pi i$ even though the circle is contained in the annulus. Note, however, that the interior of $|z| = 1$ is not contained in the annulus, so that Cauchy's theorem is not applicable. For a simply connected domain, it is true that the integral around every simple closed contour in the domain is zero.

In view of Morera's theorem, we can say that a necessary and sufficient condition for a *continuous* function to be analytic in a simply connected domain is that the integral be independent of the path of integration. At first glance, it appears that Morera's theorem is useless for proving a function to be analytic, in as much as it is not possible to test all simple closed contours. However, the proof of the next theorem should dispel any doubts as to the utility of Morera's theorem.

Theorem 8.16. *Let $\{f_n(z)\}$ be a sequence of analytic functions converging uniformly to a function $f(z)$ on all compact subsets of a domain D. Then $f(z)$ is analytic in D.*

Proof. It suffices to show that $f(z)$ is analytic at an arbitrary point z_0 in D. Construct a neighborhood D' of z_0 contained in D. By Theorem 6.26, $f(z)$ must be continuous at all points in D'. According to Theorem 8.11,

$$\lim_{n\to\infty} \int_C f_n(z)\, dz = \int_C f(z)\, dz \tag{8.15}$$

along every simple closed contour C contained in D'. Since $f_n(z)$ is analytic in D',

$$\int_C f_n(z)\, dz = 0$$

for each n and each simple closed contour C. In view of (8.15),

$$\int_C f(z)\, dz = 0.$$

By Morera's theorem, $f(z)$ must be analytic in D'. In particular, $f(z)$ is analytic at z_0. This completes the proof. ■

Remark 8.17. Requiring uniform convergence only on compact subsets, rather than on the whole domain, will give us the needed flexibility to deal with certain questions in later chapters. ●

Rewriting Theorem 8.16 in terms of series, we have the following: If $\{f_n(z)\}$ is a sequence of analytic functions, and $\sum_{k=0}^{\infty} f_k(z)$ converges uniformly to $f(z)$ on compact subsets of D, then $f(z)$ is analytic in D.

This follows on noting that, for each simple closed contour C contained in D,

$$\sum_{k=0}^{\infty} \left(\int_C f_k(z)\, dz \right) = \int_C \left(\sum_{k=0}^{\infty} f_k(z) \right) dz = \int_C f(z)\, dz = 0.$$

Corollary 8.18. *Suppose $\{f_n(z)\}$ is a sequence of functions analytic in a domain D, and that $f(z) = \sum_{n=0}^{\infty} f_n(z)$, the series being uniformly convergent on all compact subsets of D. Then for all z in D,*

$$f'(z) = \sum_{n=0}^{\infty} f_n'(z).$$

Proof. According to Theorem 8.16, $f(z)$ is analytic in D. Hence, by Theorem 8.3,

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

where C is any simple closed contour in D', a neighborhood of z contained in D. Also note that, for each n,

$$f_n'(z) = \frac{1}{2\pi i} \int_C \frac{f_n(\zeta)}{(\zeta - z)^2}\, d\zeta.$$

Since $\sum_{n=0}^{\infty} [f_n(\zeta)/(\zeta - z)^2]$ converges uniformly to $f(\zeta)/(\zeta - z)^2$ for ζ on C, we have

$$\begin{aligned} f'(z) &= \frac{1}{2\pi i}\int_C \frac{f(\zeta)}{(\zeta - z)^2}\, d\zeta \\ &= \frac{1}{2\pi i}\int_C \left(\sum_{n=0}^{\infty} \frac{f_n(\zeta)}{(\zeta - z)^2}\right) d\zeta \\ &= \sum_{n=0}^{\infty}\left(\frac{1}{2\pi i}\int_C \frac{f_n(\zeta)}{(\zeta - z)^2} d\zeta\right) \\ &= \sum_{n=0}^{\infty} f_n'(z). \end{aligned}$$

More generally, for each integer k,

$$f^{(k)}(z) = \frac{k!}{2\pi i}\int_C \frac{f(\zeta)}{(\zeta - z)^{k+1}}\, d\zeta \quad \text{and} \quad f_n^{(k)}(z) = \frac{k!}{2\pi i}\int_C \frac{f_n(\zeta)}{(\zeta - z)^{k+1}}\, d\zeta.$$

Since $\sum_{n=0}^{\infty}[f_n(\zeta)/(\zeta - z)^{k+1}]$ converges uniformly to $f(\zeta)/(\zeta - z)^{k+1}$ for ζ on C, we can conclude, as above, that

$$f^{(k)}(z) = \sum_{n=0}^{\infty} f_n^{(k)}(z) \quad \text{for all } z \text{ in } D. \qquad \blacksquare$$

Example 8.19. For example, $\sum_{n=1}^{\infty} 3^{-n}\sin(nz)$ represents an analytic function in the strip $|\operatorname{Im} z| < \ln 3$. Indeed, as

$$\begin{aligned} |3^{-n}\sin(nz)| &= 3^{-n}\left|\frac{e^{inz} - e^{-inz}}{2}\right| \\ &\leq \frac{3^{-n}}{2}\left(2e^{n|\operatorname{Im} z|}\right) \\ &= e^{-n(\ln 3 - |\operatorname{Im} z|)}, \end{aligned}$$

the Weierstrass M-test shows that $\sum_{n=1}^{\infty} 3^{-n}\sin(nz)$ converges uniformly on each compact subset of $D = \{z : |\operatorname{Im} z| < \ln 3\}$. By Corollary 8.18, the given series of functions represents an analytic function for $|\operatorname{Im} z| < \ln 3$. ●

Remark 8.20. Note a difference between real and complex series. The real series

$$f(x) = \sum_{n=1}^{\infty} \frac{\sin nx}{n^2}$$

is uniformly convergent on the real line. But a term-by-term differentiation leads to

$$f'(x) = \sum_{n=1}^{\infty} \frac{\cos nx}{n},$$

which does not converge at $x = 0$. In the complex case, a series of analytic functions uniformly convergent on compact subsets of a domain may be differentiated term-by-term to obtain the derivative of the sum. However, we cannot extend this to the boundary of the domain. For example, even though the function

$$f(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}$$

is uniformly convergent in the disk $|z| \leq 1$, term-by-term differentiation at $z = 1$ would yield

$$f'(1) = \sum_{n=1}^{\infty} \frac{1}{n},$$

which does not converge. ●

Suppose $f(z)$ is an entire function. According to Theorem 8.8, $f(z)$ has a power series representation

$$f(z) = \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n$$

valid for all z. By Theorem 6.51 or by Corollary 8.18, the derivative of the sum is the sum of the derivatives. That is,

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}$$

for all z. Now by Corollary 8.12, we may also integrate term-by-term. In other words,

$$\begin{aligned} F(z) = \int_0^z f(\zeta)\, d\zeta &= \int_0^z \left(\sum_{n=0}^{\infty} a_n \zeta^n \right) d\zeta \\ &= \sum_{n=0}^{\infty} \frac{a_n}{n+1} z^{n+1}, \end{aligned}$$

where the integral is taken along any contour joining the origin to z.

These results may be combined to obtain useful power series relationships.

Example 8.21. Let us now expand $f(z) = \sin^2 z$ in a Maclaurin series. We could, of course, take derivatives to obtain the Maclaurin expansion directly. But let us consider other methods.

Method 1. We may use (8.12) and obtain

$$\sin^2 z = \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots \right) \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots \right), \quad z \in \mathbb{C}.$$

Collecting terms and arranging in ascending order, we obtain

$$\sin^2 z = z^2 - \frac{1}{3!} z^4 + \left(\frac{2}{5!} + \frac{1}{(3!)^2} \right) z^6 + \cdots, \quad z \in \mathbb{C}.$$

In this form, it is difficult to find the general term.

Method 2. We have

$$f'(z) = 2\sin z\cos z = \sin 2z = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)!}(2z)^{2n-1}, \quad z \in \mathbb{C}.$$

But, as f is analytic in $\mathbb{C}$ and $f(0) = 0$,

$$\begin{aligned} f(z) - f(0) = \sin^2 z &= \int_0^z f'(\zeta)\,d\zeta \\ &= \sum_{n=1}^{\infty} \left(\int_0^z \frac{(-1)^{n+1}}{(2n-1)!}(2z)^{2n-1} \right) \\ &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}2^{2n-1}}{(2n)!} z^{2n}. \end{aligned}$$

Method 3. We use the trigonometric identity $\sin^2 z = (1 - \cos 2z)/2$, and so by (8.13), we obtain

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

Example 8.22. To expand $f(z) = \operatorname{Log}(1+z)$ in a Maclaurin series valid for $|z| < 1$, we rely upon the geometric series

$$f'(z) = \frac{1}{1+z} = 1 - z + z^2 - z^3 + \cdots = \sum_{n=0}^{\infty} (-1)^n z^n.$$

Hence, as f is analytic for $|z| < 1$, with $f(0) = \operatorname{Log} 1 = 0$,

$$\begin{aligned} f(z) - f(0) = \operatorname{Log}(1+z) &= \int_0^z f'(\zeta)\,d\zeta = \sum_{n=0}^{\infty} \int_0^z (-1)^n \zeta^n\,d\zeta \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n z^{n+1}}{n+1} \\ &= z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots, \quad |z| < 1. \end{aligned}$$

Note that we have chosen the principal branch so that $\operatorname{Log} 1 = 0$. More generally, when $\log 1 = 2k\pi i$, we have

$$f(z) - f(0) = \log(1+z) - 2k\pi i = \int_0^z f'(\zeta)\,d\zeta = \sum_{n=0}^{\infty} \frac{(-1)^n z^{n+1}}{n+1},$$

and so

$$\log(1+z) = 2k\pi i + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} z^n. \qquad \bullet$$

Questions 8.23.

1. If $f(z_0) = (1/2\pi i)\int_C (f(z)/(z-z_0))\,dz$ for all points z_0 inside C, is $f(z)$ analytic inside C?
2. Suppose that f is analytic inside and on a simple closed contour C, and z_0 lies outside C. What is the value of $\int_C (f(z)/(z-z_0))\,dz$?
3. Suppose that f is analytic inside and on a simple closed contour C. Does $\int_C (f'(z)/(z-z_0))\,dz = \int_C (f(z)/(z-z_0)^2)\,dz$ for all z_0 not on C?
4. If the derivatives of all orders for two different functions agree at one point, how do the two functions compare?
5. If $\lim_{n\to\infty}\int_C f_n(z)\,dz = \int_C (\lim_{n\to\infty} f_n(z))\,dz$, does $\{f_n(z)\}$ converge uniformly on C?
6. If a sequence of functions converges uniformly on all compact subsets of a domain, must the convergence be uniform throughout the domain?
7. If $\{f_n(z)\}$ is a sequence of functions analytic in a domain D, and $\{f_n(z)\}$ converges to $f(z)$ in D, is $f(z)$ analytic in D?
8. Suppose $\{f_n(z)\}$ converges uniformly to an analytic function. What can we say about the functions $\{f_n(z)\}$?
9. Does $\int_{|z|=1} z^{-3}e^{iz}\,dz = 0$? Does $\int_{|z|=1} z^{-3}\sin z\,dz = 0$?
10. If $f(z)$ is continuous inside and on a simple closed contour C, and $\int_C f(z)\,dz = 0$, is $f(z)$ analytic inside C?
11. If $f^{(k)}(z) = \sum_{n=0}^{\infty} f_n^{(k)}(z)$, what can we say about $f^{(k-1)}(z)$?
12. If $f(z) = \sum_{n=0}^{\infty} a_n z^n$, can $f'(z_0)$ exist and not equal $\sum_{n=1}^{\infty} n a_n z_0^{n-1}$?
13. Does $\int_C (z-a)^{-1}(z-b)^{-1}\,dz = 0$ for every simple closed contour C not passing through a, b?

Exercises 8.24.

1. Prove Cauchy's integral formula for multiply connected regions.
2. (a) If $P(z)$ is a polynomial of degree n, prove that

$$\int_{|z|=2} \frac{P(z)}{(z-1)^{n+2}}\,dz = 0.$$

   (b) If n and m are positive integers, show that

$$\int_C \frac{(n-1)!e^z}{(z-z_0)^n}\,dz = \int_C \frac{(m-1)!e^z}{(z-z_0)^m}\,dz$$

   along any contour containing z_0.
3. Evaluate the following integrals, where C is the circle $|z| = 3$.

   (a) $\displaystyle\int_C \frac{e^z}{z-2}\,dz$ (b) $\displaystyle\int_C \frac{e^{z^2}}{z-2}\,dz$ (c) $\displaystyle\int_C \frac{e^{z^2}}{(z-2)^2}\,dz$

   (d) $\displaystyle\int_C \frac{e^z \sin z}{(z-2)^2}\,dz$ (e) $\displaystyle\int_C \frac{e^{-z}\cos z}{(z-2)^3}\,dz$ (f) $\displaystyle\int_C \frac{3z^4+2z-6}{(z-2)^3}\,dz.$

4. Use partial fractions to evaluate the following integrals.

   (a) $\displaystyle\int_{|z|=2} \frac{1}{z^2-1}\,dz$ (b) $\displaystyle\int_{|z|=2} \frac{1}{z^2+1}\,dz$

   (c) $\displaystyle\int_{|z|=2} \frac{1}{z^4-1}\,dz$ (d) $\displaystyle\int_{|z|=3} \frac{z^3+3z-1}{(z-1)(z+2)}\,dz.$

5. If $C = \{z : |z| = r\}$ $(r \neq 1)$, then show that

$$\int_C \frac{dz}{1+z^2} = \begin{cases} -2\arctan r & \text{if } 0 < r < 1 \\ \pi - 2\arctan r & \text{if } 1 < r. \end{cases}$$

6. Use the Cauchy integral formula to evaluate the following integrals.

   (a) $\displaystyle\int_{|z|=1} (\operatorname{Re} z)^2\,dz$ (b) $\displaystyle\int_{|z-1|=1} (\overline{z})^2\,dz$ (c) $\displaystyle\int_{|z-1|=1} (\operatorname{Im} z)^2\,dz.$

7. Evaluate the integral $\int_C z/((16-z^2)(z+i))\,dz$, where C is the circle

   (a) $|z| = 2$ (b) $|z-4| = 2$ (c) $|z+4| = 2$

   (d) $|z| = \frac{1}{2}$ (e) $|z| = 5$.

8. Let $\gamma : [0, 4\pi] \to \mathbb{C}$ be given by

$$\gamma : \gamma(t) = \begin{cases} 3te^{it} & \text{if } 0 \leq t \leq 2\pi \\ 10\pi - 2t & \text{if } 2\pi \leq t \leq 4\pi. \end{cases}$$

   Evaluate the integral $\int_\gamma (z^2+\pi^2)^{-1}\,dz$.

9. Find the first five coefficients in the Maclaurin expansion for

   (a) $e^z \sin z$ (b) $\dfrac{1}{\cos z}$ (c) e^{z+z^2} (d) $e^{z/(1-z)}$.

10. Suppose $f(z)$ and $g(z)$ are analytic at z_0 with

$$f(z_0) = f'(z_0) = \cdots = f^{(n-1)}(z_0) = 0,$$
$$g(z_0) = g'(z_0) = \cdots = g^{(n-1)}(z_0) = 0.$$

   If $g^{(n)}(z_0) \neq 0$, show that

$$\lim_{z \to z_0} \frac{f(z)}{g(z)} = \frac{f^{(n)}(z_0)}{g^{(n)}(z_0)}.$$

   This is a generalization of Theorem 5.26.

11. Evaluate the following limits, using either the previous exercise or Theorem 8.8.

   (a) $\displaystyle\lim_{z\to 0} \frac{e^z-1-z}{z^2}$ (b) $\displaystyle\lim_{z\to 0} \frac{\sin z}{z-z^3}$

   (c) $\displaystyle\lim_{z\to 0} \frac{\sin z}{e^z-1}$ (d) $\displaystyle\lim_{z\to 0} \frac{\sin z - z}{\cos z - 1}.$

12. If $f(z)$ is analytic at $z = z_0$, and $|f^{(n)}(z_0)| \leq n^k$ for each n (k fixed), show that $f(z)$ is actually an entire function.
13. Show that the following series represent analytic functions in the given domain, and find their derivatives.

$$\text{(a)} \sum_{n=1}^{\infty} \frac{1}{n^z} \quad (\operatorname{Re} z > 1) \quad \text{(b)} \sum_{n=1}^{\infty} \frac{1}{z^2 - n^2} \quad (z \neq \pm 1, \pm 2, \dots).$$

14. Suppose $f(z)$ and $g(z)$ are analytic in a simply connected domain D. Prove that

$$\int_{z_0}^{z_1} f(z)g'(z)\, dz = f(z_1)g(z_1) - f(z_0)g(z_0) - \int_{z_0}^{z_1} g(z)f'(z)\, dz,$$

where the path of integration is any contour from z_0 to z_1 that lies in D.
15. Suppose $f(z)$ is continuous, but not necessarily analytic, on a contour C. Show that the function

$$F(z) = \int_C \frac{f(\zeta)}{\zeta - z}\, d\zeta$$

is analytic at each z not on C, with

$$F'(z) = \int_C \frac{f(\zeta)}{(\zeta - z)^2}\, d\zeta.$$

16. Choose a specific determination, find Maclaurin expansions for the following functions, and state the region for which the expansion is valid.

(a) $(1+z)^\alpha$ $(0 < \alpha < 1)$ (b) $\tan^{-1}(z)$ (c) $\sin^{-1} z$.

8.2 Cauchy's Inequality and Applications

In elementary calculus, we often deduce information about a function based on the behavior of its derivative. For example, if the derivative is positive, negative, or zero on an interval the function is, respectively, increasing, decreasing or constant on that interval. In complex analysis, the opposite is also true in the following sense. The behavior of an analytic function is used to estimate the behavior of its derivatives. More precisely, we see that if an analytic function f is bounded in a neighborhood of a point z_0, then the derivatives of f cannot be arbitrarily large at z_0.

Theorem 8.25. (Cauchy's Inequality) *Suppose $f(z)$ is analytic inside and on the circle C having center at z_0 and radius r. If $|f(z)| \leq M$ on C, then*

$$|f^{(n)}(z_0)| \leq Mn!/r^n, \quad n = 1, 2, \dots .$$

Proof. By the generalized Cauchy integral formula,

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz.$$

Hence, by the hypothesis,

$$\begin{aligned} |f^{(n)}(z_0)| &\leq \frac{n!}{2\pi} \int_C \frac{|f(z)|}{|z-z_0|^{n+1}}\, |dz| \\ &\leq \frac{n!M}{2\pi r^{n+1}} \int_C |dz| \\ &= \frac{n!M}{2\pi r^{n+1}} (2\pi r) = \frac{n!M}{r^n}. \end{aligned}$$ ∎

As a comparison with real analysis, we consider $f(x) = \sin(1/x)$ for $x > 0$. Then f is differentiable for $x > 0$ and $|f(x)| \leq 1$ for all $x > 0$. However,

$$f'(x) = -\frac{1}{x^2} \cos(1/x)$$

which is clearly not bounded because, for example, for each $n \in \mathbb{N}$

$$\left| f'\left(\frac{1}{n\pi}\right) \right| = n^2\pi^2.$$

Here is a surprising application of Cauchy's inequality.

Corollary 8.26. *Suppose that f is analytic for $|z-z_0| < r$. If $|f(z)-b| \leq M$, then $|f'(z_0)| \leq M/r$.*

Proof. Take $0 < \delta < r$. Then $g(z) = f(z) - b$ is analytic and maps the disk $|z-z_0| \leq \delta$ into $|w| \leq M$, so by Cauchy's inequality

$$|f'(z_0)| = |g'(z_0)| \leq M/\delta.$$

Letting $\delta \to r$ proves the result. ∎

In particular, if $f : \Delta \to \Delta$ is analytic, then $|f'(0)| \leq 1$ which is a basic version of the "Schwarz lemma" which will be discussed in the next section.

Recall the functions that are analytic in $\mathbb{C}$ are called entire functions. We know that $f : \mathbb{C} \to \mathbb{C}$ is entire if and only if $f(z)$ has the series expansion $f(z) = \sum_{n=0}^{\infty} a_n z^n$ with infinite radius of convergence. An entire function that is not a polynomial is said to be a *transcendental* entire function. The functions e^z, $\sin z$, $\cos z$, etc. are transcendental functions.

Cauchy's inequality enables us to obtain results in complex analysis which have no real variable counterpart. For example, we have

Theorem 8.27. (Liouville's Theorem) *A bounded entire function must be a constant.*

Proof. Suppose $f(z)$ is entire with $|f(z)| \leq M$ for all z. Given any complex number z_0, we have by Cauchy's inequality, $|f'(z_0)| \leq M/r$ for every positive number r. Letting $r \to \infty$, we deduce that $f'(z_0) = 0$. Since z_0 was arbitrary, $f'(z) = 0$ for all z, and hence f is constant by Theorem 5.9. ■

Remark 8.28. There is, of course, no real variable analog to Liouville's theorem. The function $f(x) = \sin x$ is a nonconstant, everywhere real differentiable function on $\mathbb{R}$ and $|f(x)| \leq 1$ on $\mathbb{R}$. On the other hand, we have already seen that each of $|\sin z|$ and $|\cos z|$ approaches to ∞ as $y \to \pm\infty$ for any fixed x. Likewise, $|\sinh z|$ and $|\cosh z|$ are unbounded entire functions. Again, by Liouville's theorem, each of these hyperbolic functions must be unbounded, because each is a nonconstant entire function. ●

Corollary 8.29. *Every $f : \mathbb{C} \to \Delta$ which is analytic is constant. In particular, there exists no bijective mapping of the unit disk Δ onto $\mathbb{C}$.*

Example 8.30. The method used in the proof of Liouville's theorem helps us to characterize all those entire functions f such that

$$|f(z)| \leq |z|^4/\ln|z| \quad \text{for } |z| > 1.$$

To do this, we choose R with $R > 1$. Then for $|z| = R > 1$,

$$|f(z)| \leq \frac{|z|^4}{\ln|z|} = \frac{R^4}{\ln R}$$

and so, by the Cauchy integral formula and the M-L Inequality

$$|a_n| = \left| \frac{1}{2\pi} \int_{|z|=R} \frac{f(z)}{z^{n+1}}\, dz \right| \leq \frac{1}{R^{n-4}\ln R} \to 0$$

as $R \to \infty$ and $n - 4 \geq 0$. Thus, $a_n = 0$ for $n \geq 4$ and hence, $f(z)$ is a polynomial of degree at most 3. ●

Liouville's theorem says that for a nonconstant entire function $f(z)$, there is a sequence of points $\{z_n\}$ such that $f(z_n) \to \infty$. This result can be sharpened.

Theorem 8.31. (Generalized version of Liouville's Theorem) *A nonconstant entire function comes arbitrarily close to every complex number.*

Proof. Suppose that $f(z)$ is entire and that there exists a complex number a such that $|f(z) - a| \geq \epsilon$ for all z. Then the function

$$g(z) = 1/[f(z) - a]$$

is entire and

$$|g(z)| = \frac{1}{|f(z) - a|} \leq 1/\epsilon.$$

By Liouville's theorem, $g(z)$ is a constant. Hence, $f(z) = 1/g(z) + a$ must also be a constant. ■

Corollary 8.32. *Suppose $f(z)$ is a nonconstant entire function. Given any complex number a, there exists a sequence $\{z_n\}$ such that $f(z_n) \to a$.*

Proof. According to Theorem 8.31, for each n we can find a point z_n such that $f(z_n) \in N(a; 1/n)$. Since $1/n \to 0$, it follows that $f(z_n) \to a$. ■

Remark 8.33. Though a nonconstant entire function come arbitrarily close to every complex value, it does not necessarily assume every complex value. For example, $f(z) = e^z$ is never equal to zero. However, $f(-n) = e^{-n} \to 0$ as $n \to \infty$. The fact that e^z assumes every other complex value can be viewed as a special case of Picard's Theorem. ●

Theorem 8.34. (Picard's Theorem) *A nonconstant entire function assumes each complex value, with one possible exception.*

For a proof of Picard's theorem, see DePree and Oehring [DO] and Ponnusamy [P1].

Our next theorem gives an estimate on the "rate of growth" of entire functions.

Theorem 8.35. *Suppose that $f(z)$ is an entire function and that $|f(z)| \leq Mr^\lambda$ ($|z| = r \geq r_0$) for some nonnegative real number λ. Then $f(z)$ is a polynomial of degree at most λ.*

Proof. Let

$$f(z) = \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n.$$

By Cauchy's inequality, on the circle $|z| = r$ we have

$$|a_n| = \frac{|f^{(n)}(0)|}{n!} \leq \frac{Mr^\lambda}{r^n} = \frac{M}{r^{n-\lambda}}.$$

Letting $r \to \infty$, we see that $a_n = 0$ whenever $n > \lambda$. Hence, $f(z)$ is a polynomial of degree no more than λ. ■

Example 8.36. If f is entire such that $|f(z)| \leq a + b|z|$ for some $a \geq 0$, $b > 0$, then f is either constant or a first degree polynomial. Let us use the Cauchy inequality to provide a proof. We let z_0 be an arbitrary point of $\mathbb{C}$. Then for $|z - z_0| \leq R$, one has

$$\begin{aligned} |f(z)| &\leq a + b|z| = a + b|z - z_0 + z_0| \\ &\leq a + b(R + |z_0|). \end{aligned}$$

By Cauchy's inequality, with $M = a + b(R + |z_0|)$

$$|a_2| = \left| \frac{f''(z_0)}{2!} \right| \leq \frac{a + b(R + |z_0|)}{R^2} \to 0 \quad \text{as} \quad R \to \infty$$

so that $f''(z_0) = 0$. Since z_0 is arbitrary, $f''(z) = 0$ on $\mathbb{C}$, and therefore $f'(z)$ is a constant, say c. Thus,

$$f(z) - f(0) = \int_0^z f'(\zeta)\, d\zeta = cz$$

and hence, $f(z) = f(0) + cz$. ●

Remark 8.37. When $\lambda = 0$, Theorem 8.35 reduces to Liouville's theorem. Choosing $0 < \lambda < 1$ shows that we need not assume that $f(z)$ is bounded (only that it grows at a sufficiently slow rate) in order to deduce that $f(z)$ must be constant. ●

Set $M(r, f) = \max_{|z|=r} |f(z)|$. Theorem 8.35 says that, for a transcendental entire function $f(z)$, the function $M(r, f)$ grows faster than any power of r. This does not mean that $f(z) \to \infty$ along every path to ∞. For instance, we have $M(r, e^z) = e^r$, but $e^z \to 0$ as $z \to \infty$ along the negative real axis. Polynomials are somewhat different. The growth of a polynomial is determined by its degree – a fact that has been used in almost every proof of the fundamental theorem of algebra.

Theorem 8.38. (Growth Lemma) *Suppose $P(z) = a_0 + a_1 z + \cdots + a_n z^n$, $a_n \neq 0$. Then there exists a sufficiently large r such that*

$$\frac{|a_n|}{2}|z|^n \leq |P(z)| \leq \frac{3|a_n|}{2}|z|^n, \quad \textit{for all } z \in \mathbb{C} \textit{ with } |z| \geq r.$$

Proof. For $z \neq 0$, we have

$$P(z) = z^n \left(a_n + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \cdots + \frac{a_0}{z^n} \right).$$

By the triangle inequality,

$$\begin{aligned} |z|^n \left(|a_n| - \left| \frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n} \right| \right) &\leq |P(z)| \\ &\leq |z|^n \left(|a_n| + \left| \frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n} \right| \right). \end{aligned}$$

For $|z| > 1$ and $n > k$, we have $|z|^n > |z|^k$ and so

$$\left| \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \cdots + \frac{a_0}{z^n} \right| \leq \frac{|a_{n-1}| + |a_{n-2}| + \cdots + |a_0|}{|z|} := \frac{K}{|z|}.$$

Hence

$$|z|^n \left(|a_n| - \frac{K}{|z|} \right) \leq |P(z)| \leq |z|^n \left(|a_n| + \frac{K}{|z|} \right) \quad (|z| \geq 1).$$

The result now follows when $K/|z| < |a_n|/2$, i.e., when

$$|z| \geq \max\{1, 2K/|a_n|\}.$$ ■

Theorems 8.35 and 8.38 provide a comparison between the growth of polynomial and transcendental functions. If $P(z)$ is a polynomial and $f(z)$ is a transcendental function, then $M(r,P)/M(r,f) \to 0$ as $r \to \infty$. That is, along its "best path", a transcendental function approaches infinity more rapidly than does a polynomial. However, $P(z) \to \infty$ as $z \to \infty$ along any path. We shall show in Chapter 9 that no transcendental function has this property. Loosely speaking, a transcendental function grow more rapidly than a polynomial, whereas a polynomial grows more consistently than a transcendental function.

An important property of polynomials is stated in Theorem 8.39.

Theorem 8.39. (The Fundamental Theorem of Algebra) *Every nonconstant polynomial has at least one zero.*

Proof. Suppose $P(z) = a_0 + a_1 z + \cdots a_n z^n$, $a_n \neq 0$. If $P(z)$ never vanishes, then $1/P(z)$ is entire. By Theorem 8.38, $1/P(z) \to 0$ as $z \to \infty$. Thus $|1/P(z)| < 1$ for $|z| \geq R$. But $1/P(z)$ is continuous (hence bounded) on the compact set $|z| \leq R$. Therefore, $1/P(z)$ is bounded in the whole plane and, by Liouville's theorem, must be a constant. This implies that $P(z)$ is also a constant, contradicting our assumption. ∎

Corollary 8.40. *Every polynomial of degree n has exactly n (not necessarily distinct) zeros.*

Proof. The fundamental theorem shows the existence of at least one zero r_1. The expression $z - r_1$ may be factored out, leaving a polynomial of degree $n-1$. Reapplying the theorem to this new polynomial, we obtain another zero. This process can be repeated n times. ∎

Corollary 8.41. *Every polynomial of degree n assumes each complex number exactly n times.*

Proof. If $P(z)$ is a polynomial of degree n, then

$$Q(z) = P(z) - a$$

is also a polynomial of degree n. By Corollary 8.40, $Q(z)$ has n zeros. But the zeros of $Q(z)$ are the "a" points of $P(z)$. ∎

Remark 8.42. Corollary 8.41 provides a more complete solution, in the case of polynomials, than does Theorem 8.31. A polynomial of degree n not only comes arbitrarily close to every complex value, it actually takes on every value n times. But the existence of n roots tells nothing about their location. In Section 9.4, we shall develop a method for approximating the location of zeros for some analytic functions. ●

Example 8.43. We wish to show that if all the zeros of a polynomial $P(z)$ have positive real parts, then so do the zeros of its derivative $P'(z)$.

To do this, let $P(z)$ be a polynomial of degree n and have zeros at z_k (not necessarily distinct) with $\mathrm{Re}\, z_k > 0$ for each $k = 1, 2, \dots, n$. By the fundamental theorem of algebra we can express $P(z)$ as

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

where c is a nonzero constant. For $z \neq z_k$, we have

$$\frac{P'(z)}{P(z)} = \frac{1}{z - z_1} + \frac{1}{z - z_2} + \cdots + \frac{1}{z - z_n}.$$

We claim that $P'(z) \neq 0$ whenever $\mathrm{Re}\, z \leq 0$. If $\mathrm{Re}\, z \leq 0$, as $\mathrm{Re}\, z_k > 0$, we have $\mathrm{Re}\,(z - z_k) < 0$. Thus, for each $k = 1, 2, \dots, n$, we see that

$$\mathrm{Re}\left(\frac{1}{z - z_k}\right) < 0$$

and therefore,

$$\mathrm{Re}\left(\frac{P'(z)}{P(z)}\right) < 0$$

which gives that $P'(z) \neq 0$ whenever $\mathrm{Re}\, z \leq 0$. Hence, the zeros of $P'(z)$ must have its real parts positive. ●

We now see how the behavior of an analytic function at a sequence of points influences its behavior elsewhere (see also Theorem 6.56).

Theorem 8.44. *Suppose $f(z)$ is analytic in the disk $|z - z_0| < R$, and that $\{z_n\}_{n \geq 1}$ is a sequence of distinct points converging to z_0. If $f(z_n) = 0$ for each $n \in \mathbb{N}$, then $f(z) \equiv 0$ everywhere in $|z - z_0| < R$.*

Proof. We have

$$f(z) = a_0 + \sum_{k=1}^{\infty} a_k (z - z_0)^k \quad (|z - z_0| < R). \tag{8.16}$$

Since $f(z)$ is continuous at z_0, it follows that $f(z_n) \to f(z_0)$ as $z_n \to z_0$. Therefore,

$$\lim_{n \to \infty} f(z_n) = 0 = f(z_0) = a_0.$$

Hence $f(z)$ has no constant term, and we may write

$$f(z) = (z - z_0)\left(a_1 + \sum_{k=2}^{\infty} a_k (z - z_0)^{k-1}\right).$$

Setting $z = z_n$ and dividing by $z_n - z_0$ leads to

$$\lim_{n\to\infty} \frac{f(z_n)}{z_n - z_0} = 0 = a_1.$$

In like manner,

$$f(z) = (z - z_0)^2 \left(a_2 + \sum_{k=3}^{\infty} a_k (z - z_0)^{k-2} \right)$$

so that

$$\lim_{n\to\infty} \frac{f(z_n)}{(z_n - z_0)^2} = 0 = a_2.$$

An induction shows that $a_k = 0$ for each k, and the result follows. ■

Corollary 8.45. (Zeros are isolated) *Suppose $f(z)$ is analytic at a point $z = z_0$. Then either $f(z) \equiv 0$ in some neighborhood of z_0, or there exists a real number r such that $f(z) \neq 0$ in the punctured disk $0 < |z - z_0| \leq r$.*

Proof. Assume that no such r exists. Then in each punctured disk $0 < |z - z_0| \leq 1/n$, there exists a point z_n such that $f(z_n) = 0$. Since $z_n \to z_0$, an application of Theorem 8.44 shows that $f(z)$ must be identically zero in some neighborhood of z_0. ■

Remark 8.46. There is no real variable analog to this corollary. The function

$$f(x) = \begin{cases} x^2 \sin \dfrac{\pi}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is differentiable for all real x, with $f(1/n) = 0$ for each n. However, $f(x) \neq 0$ in any neighborhood of the origin. ●

We now generalize the previous theorem to arbitrary domains.

Theorem 8.47. (Uniqueness Theorem) *Suppose $f(z)$ is analytic in a domain D, and that $\{z_n\}$ is a sequence of distinct points converging to a point z_0 in D. If $f(z_n) = 0$ for each n, then $f(z) \equiv 0$ throughout D.*

Proof. Consider the following two disjoint sets:

$$A = \{a \in D : \ f(z) \equiv 0 \ \text{ in some neighborhood of } a\}$$
$$B = \{a \in D : \ f(z) \neq 0 \ \text{ for all } z \text{ in some deleted neighborhood of } a\}.$$

By Corollary 8.45, every point in D is either in A or in B. It may easily be verified that both A and B are open sets. Since the domain $D = A \cup B$ is connected, either A or B must be the empty set. By Theorem 8.44, the point z_0 is in A. Therefore, $B = \emptyset$. This means that $A = D$, and $f(z) \equiv 0$ in D. ■

As a consequence of Theorem 8.47, we have the following result which is also referred to as the uniqueness/identity principle.

Theorem 8.48. (Identity Theorem) *Suppose* $\{z_n\}$ *is a sequence of points having a limit point in a domain* D*. If* $f(z)$ *and* $g(z)$ *are analytic in* D*, with* $f(z_n) = g(z_n)$ *for each* n*, then* $f(z) \equiv g(z)$ *throughout* D*.*

Proof. Let z_0 denote the limit point of $\{z_n\}$. By Theorem 2.18, there exists a subsequence $\{z_{n_k}\}$ converging to z_0. Setting $h(z) = f(z) - g(z)$, we see that $h(z_{n_k}) = 0$ for all points of the sequence $\{z_{n_k}\}$. An application of Theorem 8.47 shows that $h(z) \equiv 0$ in D and the result follows. ∎

Remark 8.49. The requirement that the limit point z_0 be in the domain of analyticity is essential. The nonconstant function $f(z) = e^{1/(1-z)}$ is analytic in $|z| < 1$ but not at the point $z = 1$. For $z_n = 1 - 1/2n\pi i$, we have

$$e^{1/(1-z_n)} = e^{2n\pi i} = 1.$$

Note that $z_n \to 1$ as $n \to \infty$, a point at which $f(z)$ is not analytic. ●

Example 8.50. Suppose that f is entire. We wish to show that $f(\mathbb{R}) \subseteq \mathbb{R}$ if and only if $f^{(n)}(0)$ is real for each $n \in \mathbb{N}$. It is evident that if each $f^{(n)}(0) \in \mathbb{R}$, then $f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n$ shows that $f(\mathbb{R}) \subseteq \mathbb{R}$. To prove the converse, let $f(\mathbb{R}) \subseteq \mathbb{R}$. Then

$$g(z) = \overline{f(\overline{z})} = \sum_{n=0}^{\infty} \frac{\overline{f^{(n)}(0)}}{n!} z^n$$

is entire and $g(z) = f(z)$ for $z \in \mathbb{R}$. By the identity theorem, $g(z) = f(z)$ throughout $\mathbb{C}$. Moreover, the power series representation of an analytic function is unique, and we must have $f^{(n)}(0) = \overline{f^{(n)}(0)}$ for all n. ●

Example 8.51. Does there exist a function $f(z)$ analytic in $|z| < 1$ and satisfying

$$f\left(\frac{1}{2n}\right) = f\left(\frac{1}{2n+1}\right) = \frac{1}{2n} \quad (n = 1, 2, \dots)? \tag{8.17}$$

The function $f(z) = z$ is an analytic function and satisfies the condition $f(1/2n) = 1/2n$. By the identity theorem, that is the only such analytic function. Since

$$f\left(\frac{1}{2n+1}\right) \neq \frac{1}{2n},$$

there does not exist an analytic function that satisfies (8.17). Note, however, that we can construct a function in $|z| < 1$ that satisfies

$$f\left(\frac{1}{2n+1}\right) = \frac{1}{2n}$$

for every n. Setting $z = 1/(2n+1)$, we have $1/2n = z/(1-z)$, so that $f(z) = z/(1-z)$ satisfies the condition. ●

Remark 8.52. Cauchy's integral formula says that the behavior of an analytic function on a simple closed contour determines its behavior inside. The identity theorem tells us even more. It says that the behavior at any sequence of points, inside or on the simple closed contour, determines the behavior of the analytic function at all points of the domain. ●

Remark 8.53. We now have a simple way to prove the standard trigonometric identities. For example, the function

$$f(z) = \sin^2 z + \cos^2 z$$

is an entire function that is equal to one on the real axis. Hence $f(z) \equiv 1$ in the complex plane; that is, $\sin^2 z + \cos^2 z \equiv 1$ for all z. ●

Example 8.54. Let $S = \{1/n : n = 1, 2, \dots\}$. Then $S \subset [0, 1]$ and S has a limit point 0. Let f be an entire function and $f(z) = p(z)$ for $z = x$, $x \in S$, where p is polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k$. By the Uniqueness Theorem, since $0 \in [0, 1] \subseteq \mathbb{C}$, we have $f(z) = a_0 + a_1 z + \cdots + a_k z^k$ for all $z \in \mathbb{C}$. ●

Questions 8.55.

1. Is there a corresponding "Cauchy inequality" when the circle is replaced by a simple closed contour?
2. Does there exist an analytic function in a neighborhood of the origin such that $|f^{(n)}(0)| \geq (n!)^2$ for all $n \in \mathbb{N}$?
3. Does there exist a condition which ensures an entire function to be a polynomial?
4. Is $\cos z$ an entire nonconstant function? Must $\cos z$ be unbounded? Must $\sin z$ be unbounded?
5. Can a nonconstant entire function be bounded in a half-plane?
6. Can a real part of a nonconstant entire function be bounded?
7. Why is Theorem 8.31 a generalization of Liouville's theorem?
8. If a nonconstant function is analytic everywhere outside a disk, can the function be bounded?
9. If $f(z) = 1/z$, then it is bounded as $z \to \infty$ but is not constant. Does this contradict Liouville's theorem?
10. What can we say about entire functions that omit the value zero?
11. What is wrong in the following proof?
    Since $e^z \neq 0$, $1/e^z$ is bounded. Therefore, by Liouville's theorem $1/e^z$ is constant.
12. Suppose that f is entire such that $|f(z)| \to \infty$ as $|z| \to \infty$. Does f have at least one zero in $\mathbb{C}$? How do we compare with e^z?

    Note: Note that e^z is entire and has no zeros in $\mathbb{C}$. We observe that $\lim_{|z|\to\infty} |e^z| = \lim_{|z|\to\infty} e^x$ does not exist.

13. If two entire functions agree at infinitely many points, must they be equal?
14. If two entire functions agree on a segment of the real axis, must they agree on $\mathbb{C}$?
15. If f is entire such that either $\operatorname{Re} f$ or $\operatorname{Im} f$ is bounded, must f be constant?
16. If f is entire such that either $\operatorname{Re} f > -2006$ or $\operatorname{Im} f > -2006$, must f be constant?
17. If $f = u + iv$ is entire such that $au + bv \geq c$ for some real numbers a, b and c, must f be constant?
18. Let f be entire such that $f(1/n) = \cos(1/n)$ for all $n \in \mathbb{Z}$. Is $f(z) = \cos z$ for all $z \in \mathbb{C}$?
19. Let f be analytic in the punctured complex plane $\mathbb{C}\backslash\{0\}$ such that $f(1/n\pi) = \sin(n\pi)$ for all $n \in \mathbb{Z}$. Is $f(z) = \sin(1/z)$ for all $z \in \mathbb{C}\backslash\{0\}$?

Exercises 8.56.

1. If $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is analytic in $|z| < R$, and $f(x)$ is real when $-R < x < R$, show that a_n is real for each n. Also show that $f(\overline{z}) = \overline{f(z)}$.
2. Let $P(z) = a_0 + a_1 z + \cdots + a_n z^n$, $a_n \neq 0$. Given $\epsilon > 0$, show that there exists an R such that $r > R$ implies
$$(1-\epsilon)|a_n|r^n \leq |P(z)| \leq (1+\epsilon)|a_n|r^n \quad (|z| = r).$$
3. If all the zeros of a polynomial $P(z)$ have negative real parts, show that all the zeros of $P'(z)$ have negative real parts.
4. Suppose $f(z)$ is an entire function with $|f(z)| \leq |e^z|$ for all z. Prove that $f(z) = Ke^z$, $|K| \leq 1$.
5. Find all entire functions f for which there exists a positive constant M such that $|f(z)| \leq M|\cos z|$ for all $z \in \mathbb{C}$. How about if $\cos z$ is replaced by $\cosh z$ or $\sin z$ or $\sinh z$ respectively?
6. Suppose that $f(z)$ is an entire function such that $|f'(z)| \leq |z|$ for all $z \in \mathbb{C}$. Show that f must be of the form $f(z) = az^2 + b$ where a, b are complex constants such that $|a| \leq 1/2$. What will be the form of f if $f(z)$ is entire such that $|f^{(k)}(z)| \leq |z|$ for some fixed $k > 2$ and for all $z \in \mathbb{C}$?
7. Suppose that $f(z)$ is entire with a and b positive constants. If
$$f(z+a) = f(z+bi) = f(z)$$
for all z, show that $f(z)$ is constant.
8. Suppose that f is an entire function such that $|f(z)| \leq 10$ on $|z-2| = 3$. Find a bound for $|f^{(3)}(2)|$.
9. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be analytic for $|z| \leq 1$, and assume that $|f(z)| \leq 1$ for $|z| \leq 1$. Use Cauchy's inequality to prove

(a) $\left|\sum_{n=k}^{\infty} a_n z^n\right| \leq \frac{|z|^k}{1-|z|} \quad (|z|<1)$.

(b) If $|z_1| \leq r$, $|z_2| \leq r$, $0<r<1$, and $z_1 \neq z_2$, then

$$\left|\frac{f(z_2)-f(z_1)}{z_2-z_1}\right| \leq \frac{1}{(1-r)^2}.$$

10. If $f(z)$ and $g(z)$ are analytic in a domain D with $f(z)g(z)=0$ in D, prove that either $f(z) \equiv 0$ or $g(z) \equiv 0$ in D.
11. Suppose $f(z)$ and $g(z)$ are analytic in domain D and that

$$\frac{f'(z_n)}{f(z_n)} = \frac{g'(z_n)}{g(z_n)}$$

at a sequence of points $\{z_n\}$ converging to a point z_0 in D. Show that $f(z) = Kg(z)$ in D.
12. Give an example of a nonvanishing analytic function f in the unit disc $|z|<1$ having infinitely many zeros.
13. Prove that there is no analytic function f in the unit disk $\varDelta = \{z : |z|<1\}$ such that $f(1/n) = (-1)^n/n^2$ for $n = 2,3,4,\ldots$.
14. Let f and g be analytic in the unit disk $\varDelta$.
(a) If $f(1/n) = g(1/n)$ for $n = 2,3,\ldots,$ show that $f=g$.
(b) Show that $f(1/n) = 1/\sqrt{n}$ for each $n = 2,3,\ldots$ is not possible.
15. Suppose that f is entire and that there exists a bounded sequence of distinct real numbers $\{a_n\}_{n\geq 1}$ such that $f(a_n)$ is real for each $n \geq 1$. Show that $f(z)$ is real on $\mathbb{R}$. In addition, if $\{a_n\}$ is decreasing such that $a_n \to 0$ as $n \to \infty$, and $f(a_{2n}) = f(a_{2n+1})$ for all $n \geq 1$, then show that f is a constant.
16. In each case, exhibit a nonconstant f having the desired properties or explain why no such function exists:
(a) f is entire with $f^{(n)}(0) = 3^n$ for n even and $f^{(n)}(0) = (n-1)!$ for n odd.
(b) f is analytic in $|z|<1$ with $f(1/n) = n/(2+n)$ for $n \in \mathbb{N}$.
(c) f is analytic in $|z|<1$ such that $f(1/n) = (1+(-1)^n)/3$ for $n \in \mathbb{N}$.
(d) f is analytic in $|z|<1$ such that $f(1/n) = 2^n$ for $n \in \mathbb{N}$.
(e) f is analytic in $|z|<1$ such that $f(1/n) = 1/\sqrt{n}$ for each $n = 2,3,\ldots$.
17. The functions $\sqrt{z}$ and $\sin\sqrt{z}/\sqrt{z}$ are well defined and analytic on the cut plane $\mathbb{C}\backslash(-\infty,0]$. Show that

$$f(z) = \frac{\sin\sqrt{z}}{\sqrt{z}} = \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{(2k+1)!} z^k \quad \text{for } z \in \mathbb{C}\backslash(-\infty,0].$$

Explain why it is possible to define $f(z)$ on the cut $(-\infty,0]$ so that f is analytic on $\mathbb{C}$. What values $f(x)$ should be assigned when $x \in (-\infty,0]$? Can this procedure be applied to $g(z) = \sin\sqrt{z}$?

8.3 Maximum Modulus Theorem

Suppose $f(x)$ is a continuous real-valued function defined on the interval $[a, b]$ and that $F'(x) = f(x)$ throughout. Then the "average" value of $f(x)$ on the interval is given by

$$\frac{1}{b-a}\int_a^b f(x)\,dx = \frac{F(b)-F(a)}{b-a}.$$

Furthermore, according to the mean-value theorem, this expression is equal to $F'(\xi)$ for some ξ, $a < \xi < b$.

Our next theorem shows that for functions analytic inside and on a circle, the average of the values on the circumference is equal to the value of the function at the center of the circle.

Theorem 8.57. (Gauss's Mean-Value Theorem) *Suppose $f(z)$ is analytic in the closed disk $|z - z_0| \le r$. Then*

$$f(z_0) = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{i\theta})\,d\theta.$$

Proof. By Cauchy's integral formula,

$$f(z_0) = \frac{1}{2\pi i}\int_{|z-z_0|=r} \frac{f(z)}{z-z_0}\,dz.$$

Write this out in terms of a parameterization $z = z_0 + re^{i\theta}$ with $0 \le \theta \le 2\pi$, $dz = ire^{i\theta}d\theta$. Then

$$f(z_0) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{re^{i\theta}} ire^{i\theta}\,d\theta = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{i\theta})\,d\theta. \qquad \blacksquare$$

If $f(z)$ is a constant (say $f(z) = C$), then Gauss's theorem gives the obvious fact that

$$f(z_0) = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{i\theta})\,d\theta = C.$$

Our next theorem shows that for nonconstant functions there must be some points on the circle $|z| = r$ for which $|f(z)| > |f(z_0)|$.

Theorem 8.58. (Maximum Modulus Theorem: First Form) *If $f(z)$ is analytic in a domain D, then $|f(z)|$ cannot attain a maximum in D unless $f(z)$ is constant.*

Proof. Suppose $|f(z)|$ attains maximum at a point z_0 in D. Choose a disk $|z - z_0| \le r$ contained in D. Gauss's mean-value theorem gives

$$f(z_0) = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{i\theta})\,d\theta \tag{8.18}$$

so that

$$|f(z_0)| \le \frac{1}{2\pi}\int_0^{2\pi} |f(z_0 + re^{i\theta})|\,d\theta.$$

By assumption,

$$|f(z_0 + re^{i\theta})| \le |f(z_0)|, \tag{8.19}$$

and so

$$\frac{1}{2\pi}\int_0^{2\pi} |f(z_0 + re^{i\theta})|\,d\theta \le \frac{1}{2\pi}\int_0^{2\pi} |f(z_0)|\,d\theta = |f(z_0)|. \tag{8.20}$$

Combining (8.18) and (8.20), we have

$$|f(z_0)| = \frac{1}{2\pi}\int_0^{2\pi} |f(z_0 + re^{i\theta})|\,d\theta$$

or

$$\frac{1}{2\pi}\int_0^{2\pi} (|f(z_0)| - |f(z_0 + re^{i\theta})|)\,d\theta = 0.$$

By (8.19) the integrand is nonnegative and therefore

$$|f(z_0 + re^{i\theta})| = |f(z_0)| \quad \text{for} \;\; 0 \le \theta \le 2\pi.$$

Hence, $|f(z)| = |f(z_0)|$ for each z on $|z - z_0| = r$. Since r is arbitrary, $|f(z)| = |f(z_0)|$ for all points inside and on $|z - z_0| = r$. Recall that an analytic function with constant modulus in a domain is constant in that domain, hence $f(z)$ is constant on $|z - z_0| < r$. From the identity theorem, it follows that $f(z)$ is constant in the whole domain. Therefore, $|f(z)|$ cannot attain a maximum at a point of D unless $f(z)$ is constant. ∎

Suppose that f is analytic on a domain D and continuous on the boundary ∂D. By Theorem 8.58, $|f(z)|$ cannot attain its maximum in D unless $f(z)$ is constant. This raises the following questions:

- Does $|f(z)|$ attains its maximum on ∂D?
- Is $|f(z)| \le M$ on D when $|f(z)| \le M$ on ∂D?

In general, the answer to both questions are negative.

Theorem 8.59. (Maximum Modulus Theorem: Second Form) *If $f(z)$ is analytic in a bounded domain D and continuous on its closure D, then $|f(z)|$ attains a maximum on the boundary. Furthermore, $|f(z)|$ does not attain a maximum at an interior point unless $f(z)$ is constant.*

Proof. First, observe that $\overline{D}$ is a compact set because D is bounded. Thus, by Theorem 2.41, as $|f(z)|$ is a continuous real function on $\overline{D}$, $|f(z)|$ attains a maximum somewhere in $\overline{D}$. According to the first form of the theorem, the maximum cannot occur at an interior point. Hence the maximum must occur on the boundary. ∎

Remark 8.60. The domain need not be simply connected. For instance, the modulus of any function analytic in the open annulus $1/R < |z| < R$ and continuous on the closed annulus $1/R \leq |z| \leq R$ must attain its maximum on the boundary. The modulus of the function $f(z) = z$ attains its maximum on the outer boundary, whereas the modulus of $f(z) = 1/z$ attains the maximum on the inner boundary. ●

Theorem 8.59 is not true if D is not bounded. For example, if

$$f(z) = e^z \quad \text{and} \quad D = \{z \in \mathbb{C} :\, \mathrm{Re}\, z > 0\},$$

then ∂D is the imaginary axis and $|f(iy)| = |e^{iy}| = 1$, i.e., $f(\partial D) = \partial \Delta$. Yet $|f(z)| = e^x \to \infty$ as $z \to \infty$ along the positive real axis. Thus, the hypothesis that D is bounded is essential in Theorem 8.59.

Here is another example. Set $f(z) = e^{iz^2} = e^{i(x^2-y^2)}e^{-2xy}$ so that

$$|f(z)| = e^{-2xy}.$$

If $D = \{z \in \mathbb{C} :\, \mathrm{Re}\, z > 0,\ \mathrm{Im}\, z < 0\}$, then, for points on the boundary ∂D, either $x = 0$ or $y = 0$ and so $|f(z)| = 1$ on ∂D. However, for $y = -x$ $(x > 0)$ we have

$$|f(z)| = e^{2x^2} \to \infty \quad \text{as} \quad x \to \infty.$$

Then both examples show that the modulus of an analytic function need not attain its maximum on the boundary ∂D and the maximum of the modulus on the boundary may not be the maximum value inside the domain D unless D is bounded.

Example 8.61. Suppose that we wish to find the maximum modulus of $f(z) = 3z - 2i$ on $|z| \leq 3$. To do this, we compute

$$|f(z)|^2 = |3z - 2i|^2 = 9|z|^2 + 12\mathrm{Re}\,(iz) + 4 = 9|z|^2 - 12\mathrm{Im}\, z + 4.$$

By Theorem 8.59, $\max_{|z|\leq 3} |f(z)|$ occurs on the boundary $|z| = 3$. Therefore, on $|z| = 3$,

$$|f(z)| = \sqrt{9(3)^2 - 12\mathrm{Im}\, z + 4} = \sqrt{85 - 12\mathrm{Im}\, z}.$$

The last expression attains its maximum when $\mathrm{Im}\, z$ attains its minimum on $|z| = 3$, namely, at the point $z = -3i$. Thus,

$$\max_{|z|\leq 3} |f(z)| = \sqrt{85 - 12(-3)} = \sqrt{121} = 11.$$

Alternatively, as $|f(z)|$ attains its maximum on $|z|=3$, we consider $z=3e^{i\theta}$ and compute

$$|f(z)| = |3(3e^{i\theta}) - 2i| = \sqrt{85 - 36\sin\theta}.$$

This expression is maximum when $-\sin\theta$ is maximum, i.e., when $\theta = -\pi/2$. Thus, the maximum value of $|f(z)|$ on $|z| \le 3$ is 11. More generally, we have the following:

If $0 \neq a \in \mathbb{C}$, $0 \neq b \in \mathbb{C}$ and $f(z) = az + b$ on $|z| \le R$, then

$$\max_{|z|\le R} |az+b| = \max_{|z|\le 1} |aRz + b| = |a|R + |b|$$

and the maximum value is attained at z_0 on the boundary $|z| = 1$, where $\arg z_0 = \operatorname{Arg} b - \operatorname{Arg} a$. Clearly, this follows from

$$az + b = |a|e^{i\operatorname{Arg} a}z + |b|e^{i\operatorname{Arg} b} = e^{i\operatorname{Arg} b}\left[|a|ze^{-i(\operatorname{Arg} b - \operatorname{Arg} a)} + |b|\right].$$

Moreover, it is easy to see that $\max_{|z|\le 1} |az^n + b| \le |a| + |b|$. ●

Example 8.62. If $f(z) = z^2/(z^3 - 10)$ for $|z| \le 2$, then f is analytic inside and on $|z| = 2$. Then $|f(z)|$ attains its maximum value on $|z| = 2$. For $z = 2e^{i\theta}$, $0 \le \theta \le 2\pi$, we have

$$\begin{aligned}|f(z)| &= \frac{|z|^2}{|(2e^{i\theta})^3 - 10|}\\ &= \frac{4}{\sqrt{(8\cos 3\theta - 10)^2 + 8^2\sin^2 3\theta}}\\ &= \frac{4}{\sqrt{64 + 100 - 160\cos 3\theta}}.\end{aligned}$$

This expression is maximum when $-\cos 3\theta$ is minimum. Clearly, the minimum value is when $\cos 3\theta = 1$, i.e., when $\theta = 0,\ 2\pi/3,\ 4\pi/3,\ 2\pi$. Thus, the maximum value of $|f(z)|$ is 2. ●

As an application of the maximum modulus theorem, we reprove the fundamental theorem of algebra (see also Theorem 8.39).

Suppose $P(z) = a_0 + a_1 z + \cdots + a_n z^n$, $a_n \neq 0$, has no zeros. Then $1/P(z)$ is a nonconstant entire function with no zeros. In view of Theorem 8.38, $1/P(z) \to 0$ as $z \to \infty$. When r is large enough,

$$\left|\frac{1}{P(z)}\right| < \left|\frac{1}{P(0)}\right| = \left|\frac{1}{a_0}\right| \quad (|z| = r).$$

Thus the continuous function $1/P(z)$ does not attain a maximum on the boundary of the compact set $|z| \le r$. Hence $|1/P(z)|$ must attain a maximum at an interior point, contradicting the maximum modulus theorem. Therefore, $P(z)$ must have a zero in the disk $|z| \le r$, and the proof is complete.

It is possible for the modulus of a nonconstant analytic function to attain a minimum in a domain, and not on its boundary. To see this, consider $f(z) = z^n$ on $|z| \le r$. Then

$$0 = |f(0)| = \min_{|z|\le r} |f(z)| < \min_{|z|=r} |f(z)| = r^n,$$

and so the $\min\{|f(z)| : |z| \le r\}$ is attained at the interior point 0. The $\max\{|f(z)| : |z| \le r\} = r^n$ is attained at the boundary point r, on $|z| = r$. However, we do have the following counterpart to the maximum modulus theorem.

Theorem 8.63. (Minimum Modulus Theorem) *Suppose $f(z)$ is analytic in a domain D, and that $f(z) \ne 0$ in D. Then $|f(z)|$ cannot attain a minimum in D unless $f(z)$ is constant. If $f(z)$ is also continuous on $\overline{D}$, $\overline{D}$ compact, then $|f(z)|$ attains a minimum on the boundary.*

Proof. If $f(z) \ne 0$ in D, then $1/f(z)$ is analytic in D. The function $|f(z)|$ attains a minimum at a point z_0 in D if and only if $1/|f(z)|$ attains a maximum at z_0. The result follows now upon applying the maximum modulus theorem to $1/f(z)$. ■

Not surprisingly, the fundamental theorem of algebra can also be proved from the minimum modulus theorem, for, in view of Theorem 8.38, when r is large enough the polynomial

$$P(z) = a_0 + a_1 z + \cdots + a_n z^n \quad (a_n \ne 0),$$

satisfies the inequality

$$|P(z)| > |P(0)| = |a_0| \quad (|z| = r).$$

To preserve the validity of the minimum modulus theorem, the hypothesis $P(z) \ne 0$ in $|z| < r$ must be false. That is, $P(z)$ must have a zero in the disk $|z| < r$.

Example 8.64. Consider $f(z) = e^{z^2}/z$ on $\overline{D} = \{1 \le |z| \le 2\}$. We wish to find points where $|f(z)|$ has maximum and minimum values. To do this, we set $z = re^{i\theta}$. Then, on $|z| = r$,

$$|f(z)| = \frac{e^{r^2 \cos 2\theta}}{r}$$

and the maximum occurs when $\cos 2\theta = 1$ and the minimum occurs when $\cos 2\theta = -1$. Thus

$$\max_{|z|=r} |f(z)| = \frac{e^{r^2}}{r} \quad \text{and} \quad \min_{|z|=r} |f(z)| = \frac{e^{-r^2}}{r}.$$

Note that the maximum occurs at $\theta = 0,\ \pi$ while the minimum occurs at $\theta = \pi/2,\ 3\pi/2$. In particular,

$$\max_{z\in\overline{D}} |f(z)| = \frac{e^4}{2} \quad \text{and} \quad \min_{z\in\overline{D}} |f(z)| = \frac{1}{2e^4}. \qquad \bullet$$

If $f(z)$ is a nonconstant analytic function in $|z| \leq R$ and $|f(z)| \leq M$ on $|z| = R$, then the maximum modulus theorem says that $|f(z)| < M$ for $|z| < R$. We next develop methods to improve this bound inside the disk.

Lemma 8.65. (Schwarz's Lemma) *Suppose $f(z)$ is analytic for $|z| < R$ with $f(0) = 0$. If $|f(z)| \leq M$ in $|z| < R$, then*

$$|f'(0)| \leq M/R \quad \textit{and} \quad |f(z)| \leq M|z|/R \quad \textit{for } |z| < R,$$

with equality only for $f(z) = (M/R)e^{i\alpha}z$, α real.

Proof. Since $f(0) = 0$, we may write $f(z) = a_1 z + a_2 z^2 + a_3 z^3 + \cdots$. Define a function $g : \Delta \to \mathbb{C}$ by

$$g(z) = \begin{cases} f(z)/z & \text{if } 0 < |z| < R \\ f'(0) & \text{if } z = 0. \end{cases}$$

Then $g(z)$ is analytic for $|z| < R$. Since $|f(z)| \leq M$ for $|z| < R$, we have

$$\max_{|z|=r} |g(z)| \leq \frac{M}{r} \tag{8.21}$$

for all positive real number $r < R$. By the maximum modulus theorem applied to $g(z)$, $|g(z)| \leq M/r$ for all $|z| \leq r$, $0 < r < R$. Now, since r can come arbitrarily close to R, we have

$$|g(z)| \leq \lim_{r\to R} \frac{M}{r} = \frac{M}{R} \quad \text{for all } |z| < R.$$

This proves that

$$|f(z)| \leq \frac{M}{R}|z| \quad \text{for all } |z| < R$$

and therefore, $|f'(0)| = |g(0)| \leq M/R$.

In case either $|f'(0)| = M/R$ or $|f(b)| = (M/R)|b|$ for some b, $0 < |b| < R$, we get $|g(0)| = M/R$ or $|g(b)| = M/R$ and so, $|g(z)|$ attains a maximum in $|z| < R$. Therefore, $g(z)$ is a constant function by the maximum modulus principle and the result follows. ■

Remark 8.66. Schwarz's lemma can be phrased in terms of the function $M(r, f)$. For if $f(z)$ is analytic in $|z| < R$ with $f(0) = 0$, then $M(r, f) \leq (r/R')M(R', f)$ $(r < R' < R)$. •

Corollary 8.67. *Suppose $f(z)$ is analytic for $|z| < R$ with $f(0) = f'(0) = \cdots = f^{(n-1)}(0) = 0$. If $|f(z)| \le M$ in $|z| < R$, then*

$$|f(z)| \le (|z|/R)^n M \quad (|z| < R),$$

with equality only for $f(z) = (M/R^n)e^{i\alpha}z^n$, α real.

Proof. Write $f(z) = z^n g(z)$, and apply the maximum modulus theorem to $g(z)$. ■

If $f(0) \neq 0$, Schwarz's lemma may be modified to obtain Corollary 8.68.

Corollary 8.68. *Suppose $f(z)$ is analytic for $|z| < R$. If $|f(z)| \le M$ in $|z| < R$, then $|f(z) - f(0)| \le (2|z|/R)M$ $(|z| < R)$.*

Proof. Set $g(z) = f(z) - f(0)$. Then $g(0) = 0$ and

$$|g(z)| \le |f(z)| + |f(0)| \le 2M.$$

Applying Schwarz's lemma to $g(z)$, we obtain

$$|g(z)| = |f(z) - f(0)| \le \frac{2|z|}{R} M \quad (|z| < R).$$ ■

Remark 8.69. Corollary 8.68 supplies another proof of Liouville's theorem. Suppose $f(z)$ is entire with $|f(z)| \le M$ for all z. Given a point z_0, $|z_0| = r_0$, and an $\epsilon > 0$, choose $R > 2r_0 M/\epsilon$. Then

$$|f(z_0) - f(0)| \le \frac{2r_0 M}{R} < \epsilon.$$

Since ϵ was arbitrary, $f(z_0) = f(0)$. But z_0 was also arbitrary, so that $f(z) = f(0)$ for all z. That is, $f(z)$ is a constant. ●

Example 8.70. We wish to characterize those functions f such that $f(z)$ is analytic for $|z| \le 1$ and $|f(z)| = 1$ for $|z| = 1$. To do this, we split the proof into two parts.

Case (i): Let f have no zeros in $|z| < 1$. As $|f(z)| = 1$ for $|z| = 1$, the maximum and minimum modulus theorem shows that $1 \le |f(z)| \le 1$ for $|z| \le 1$; i.e., $|f(z)| = 1$ for $|z| \le 1$. By Theorem 5.37 (see also Corollary 9.57), $f(z) = e^{i\alpha}$, α real.

Case (ii): Let f have zeros in $|z| < 1$. Clearly, f cannot have infinitely many zeros in $|z| < 1$; otherwise a limit point would be in $|z| \le 1$ in which f is analytic and so $f(z) \equiv 0$ by the uniqueness theorem. This is a contradiction because $|f(z)| = 1$ for $|z| = 1$. Define

$$F(z) = \frac{f(z)}{\prod_{k=1}^{n} \left(\frac{z-a_k}{1-\overline{a}_k z}\right)},$$

where $a_k's$ are the finite zeros of f in $|z| < 1$. Then F is analytic in $|z| \leq 1$, $F(z) \neq 0$ in $|z| < 1$ and $|F(z)| = 1$ for $|z| = 1$. By Case (i),

$$F(z) = e^{i\alpha} \quad \text{or} \quad f(z) = e^{i\alpha} \prod_{k=1}^{n} \frac{z - a_k}{1 - \overline{a}_k z}$$

for some real α. ●

Example 8.71. Suppose that $f : \Delta \to \Delta$ is analytic such that $f(0) = 0$. Then it can be easily seen that

(i) $|f(z) + f(-z)| \leq 2|z|^2$ in $\Delta = \{z : |z| < 1\}$,
(ii) the inequality in (i) is strict, except at the origin, unless f has the form $f(z) = \epsilon z^2$ with $|\epsilon| = 1$.

To see this, we let $f(z) = \sum_{n=1}^{\infty} a_n z^n$. Then

$$F(z) = \frac{f(z) + f(-z)}{2z} = \sum_{n=1}^{\infty} a_{2n} z^{2n-1}$$

and, by Schwarz' lemma, $|f(z)| \leq |z|$. Therefore, $|F(z)| \leq 1$ and (i) follows. If the equality in (i) holds at some point in $|z| < 1$, then $|F(z)| = 1$ for all $|z| < 1$ and so $F(z) = \epsilon z$ with $|\epsilon| = 1$. This gives

$$f(z) + f(-z) = 2\sum_{n=1}^{\infty} a_{2n} z^{2n} = 2\epsilon z^2.$$

Comparing the coefficients of z^n on both sides shows that $a_2 = \epsilon$ and $a_{2n} = 0$ for each $n > 1$. Thus f has the form

$$f(z) = \epsilon z^2 + \sum_{n=1}^{\infty} a_{2n+1} z^{2n+1} = \epsilon z^2 + h(z),$$

where $h(z)$ is odd. We need to show that $h(z) = 0$. Note that,

$$1 \geq |f(z)|^2 = |\epsilon z^2 + h(z)|^2, \quad z \in \Delta,$$

and, since h is odd,

$$1 \geq |f(-z)|^2 = |\epsilon z^2 + h(-z)|^2 = |\epsilon z^2 - h(z)|^2, \quad z \in \Delta.$$

Adding the last two inequalities, we see that

$$2 \geq |\epsilon z^2 + h(z)|^2 + |\epsilon z^2 - h(z)|^2 = 2(|z|^4 + |h(z)|^2)$$

so that $|h(z)|^2 \leq 1 - |z|^4$ throughout $|z| < 1$. Note that h is analytic for $|z| < 1$ and by the maximum modulus theorem, $|h(z)|^2 \leq \epsilon$ when $|z|^4 \leq 1 - \epsilon$. Since $\epsilon > 0$ is arbitrary, we see that $h(z) = 0$.

Alternatively, as $|f(z)| < 1$ on Δ, for the proof of $h(z) = 0$, it suffices to observe that

$$\lim_{r\to 1^-} \frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|^2 d\theta = \sum_{n=1}^{\infty} |a_n|^2 \le 1.$$

Since $a_2 = \epsilon$, the remaining Taylor's coefficients of f must be zero, and so $f(z) = \epsilon z^2$. ●

Questions 8.72.

1. Suppose f is analytic in the annulus $1 \le |z| \le R$, $|f(z)| \le R^n$ for $|z| = R$ and $|f(z)| \le 1$ on $|z| = 1$. Is $|f(z)| \le |z|^n$ in the annulus?
2. Suppose $f(z)$ is analytic inside and on a simple closed contour C. Can $|f(z)|$ be constant on C without $f(z)$ being constant?
3. Suppose $f(z)$ is analytic in $|z| \le r$. Do there exist two distinct points $z_0 = re^{i\theta_0}$ and $z_1 = re^{i\theta_1}$ such that $|f(z)| \le |f(z_0)|$ and $|f(z)| \le |f(z_1)|$ for all z, $|z| < r$?
4. Can the modulus of a nonconstant function, analytic in a region that is not closed, attain a maximum? What if the region is unbounded?
5. Suppose $f(z)$ and $g(z)$ are analytic inside and on a simple closed contour C, and that $|f(z) - g(z)| = 0$ on C. Does $f(z) = g(z)$ inside C? What if $|f(z)| - |g(z)| = 0$ on C?
6. Define $m(r, f) = \min_{|z|=r} |f(z)|$. What properties does $m(r, f)$ have in common with $M(r, f)$?
7. Consider the function $f(z) = z^2 + 3z + 1$ for $|z| \le 1$. Then the triangle inequality gives that $|f(z)| \le 5$ for $|z| \le 1$. Can we conclude that $\max_{|z|\le 1} |f(z)| = 5$? What happens if $f(z) = z^2 - 3z - 1$? What happens if $f(z) = z^2 + 3iz + i$? What happens if $f(z) = z^2 - 3z + 1$?
8. Suppose f is a one-to-one analytic function from the unit disk $|z| < 1$ onto itself such that $f(0) = 0$. Does $f(z) \equiv e^{i\alpha} z$ for some real α?

Exercises 8.73.

1. If $f(z)$ is analytic and nonzero in the disk $|z - z_0| \le r$, show that
$$\log|f(z_0)| = \frac{1}{2\pi}\int_0^{2\pi} \log|f(z_0 + re^{i\theta})|\, d\theta.$$
2. Suppose the nonconstant function $f(z)$ is analytic in a domain D and continuous on its closure $\overline{D}$. If $|f(z)|$ is constant on the boundary of D, prove that $f(z)$ has a zero in D.
3. Suppose that f is analytic and bounded in the unit disk Δ and $f(\zeta) \to 0$ as $\zeta \to 1^-$ along the upper half of the circle $C : |\zeta - 1/2| = 1/2$. Then, $f(x) \to 0$ as $x \to 1^-$ along $[0, 1)$.
4. Set $M(r, f) = \max_{|z|=r} |f(z)|$ and $m(r, f) = \min_{|z|=r} |f(z)|$. Find $M(r, f)$ and $m(r, f)$ for the following entire function, and indicate all points on $|z| = r$ where the maximum and minimum occur.

(a) $f(z) = e^z$ (b) $f(z) = z^n$

(c) $f(z) = z^2 + 1$ (d) $f(z) = z^2 - z + 1$

5. Find the maximum and minimum values of
   (a) $|z(1-z)|$ on $|z| \le 1$
   (b) $|z/(z^2+9)|$ on $1 \le |z| \le 2$
   (c) $|(3-iz)^2|$ on $|z| \le 1$
   (d) $|5+2iz^2|$ on $|z| \le 1$
   (e) $\left|\dfrac{z-\alpha}{1-\overline{\alpha}z}\right|$ on $|z| \le 1$ (where α with $|\alpha| < 1$ is fixed).
6. Show that $\max_{|z|=r} |e^{z^2-iz}|$ is attained at the point ir when $r \le 1/4$ and at $re^{i\theta}$, $\theta = \sin^{-1}(1/4r)$, when $r > 1/4$.
7. Suppose f is analytic for $|z| \le 1$, $f(0) = 0$ and $|f(z)| \le 5$ for all $|z| = 1$. Can $|f'(0)| > 5$?
8. Let $P(z)$ be a polynomial, and set $f(z) = P(z)e^z$. Show that $M(r, f) \to \infty$ and $m(r, f) \to 0$ as $r \to \infty$.
9. Consider the polynomial $P(z) = a_0 + a_1 z + \cdots + a_{n-1}z^{n-1} + z^n$. Show that, for all sufficiently large values of r, there must exist points z_0, z_1 on the circle $|z| = r$ such that
$$\left|e^{P(z_0)}\right| = e^{(1/2)r^n}, \quad \left|e^{P(z_1)}\right| = e^{-(1/2)r^n}.$$
10. Suppose $f(z)$ is analytic with $|f(z)| < 1$ for $|z| < 1$. If $f(0) = 0$, show that $|f'(0)| \le 1$, with equality only when $f(z) = e^{i\alpha}z$ (α real).
11. Suppose $f(z)$ is analytic for $|z| \le 1$, and $|f(z)| \ge 1$ for $|z| \le 1$. If $f(0) = 1$, show that $f(z)$ is a constant.
12. Let $f(z)$ be analytic in $|z| < R$ with $f(0) = 0$. Prove that
$$F(z) = f(z) + f(z^2) + f(z^3) + \cdots$$
is analytic in $|z| < R$, and that
$$|F(z)| \le \frac{r}{1-r} \quad (|z| = r < R).$$
13. Does there exist an analytic function $f : \Delta \to \Delta$ such that $f(1/2) = 3/4$ and $f'(1/2) = 2/3$, where $\Delta = \{z : |z| < 1\}$?
14. Let f be analytic for $|z| \le 3$ such that $|f(z)| \le 1$ for $|z| \le 3$ and has n roots at $w_k = e^{2k\pi i/3}$ ($k = 0, 1, 2, \ldots, n-1$), the nth roots of unity. What is the maximum value of $|f(0)|$? Which functions attain a maximum?

9

Laurent Series and the Residue Theorem

In this chapter, we investigate the behavior of a function at points where the function fails to be analytic. While such functions cannot be expanded in a Taylor series, we show that a Laurent series expansion is possible. Also, we introduce the notion of isolated and non-isolated singularities and discuss different ways of characterizing isolated singularities. The complex integration machinery that was built in Chapter 7 and developed in Chapter 8 is now ready to be utilized in order to evaluate definite integrals of real-valued functions.

9.1 Laurent Series

We know that a function $f(z)$, analytic at a point z_0, has a power series representation $f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n$ that is valid in some neighborhood of z_0. In this section, we will characterize expressions of the form

$$\sum_{n=-\infty}^{\infty} a_n(z-z_0)^n.$$

To this end, observe that the series

$$\sum_{n=1}^{\infty} b_n(z-z_0)^{-n}$$

may be viewed as a power series in the variable $1/(z-z_0)$. If R is its radius of convergence, then the series converges absolutely for

$$1/|z-z_0| < R, \quad \text{that is, for } |z-z_0| > R_1 = 1/R.$$

Thus the series $\sum_{n=1}^{\infty} b_n(z-z_0)^{-n}$ represents an analytic function, $f_1(z)$, outside the circle $|z-z_0| = R_1$. Suppose also that the series

$$\sum_{n=0}^{\infty} a_n(z-z_0)^n$$

has radius of convergence R_2. Then, $f_2(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n$ is analytic for $|z-z_0| < R_2$. If $R_2 > R_1$, then $f_1(z)$ and $f_2(z)$ are both analytic in the annulus $R_1 < |z-z_0| < R_2$. Hence the function

$$f(z) = f_1(z) + f_2(z) = \sum_{n=1}^{\infty} b_n(z-z_0)^{-n} + \sum_{n=0}^{\infty} a_n(z-z_0)^n$$

is analytic for all z in the annulus $R_1 < |z-z_0| < R_2$. Setting $a_{-n} = b_n$, we may rewrite the above expression as

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n. \tag{9.1}$$

An expression of the form (9.1) is called a *Laurent series* about the point z_0. We remark that if $R_1 > R_2$, then the Laurent series $\sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$ diverges everywhere, and if $R_1 = R_2$, it diverges everywhere except possibly at points where $|z-z_0| = R_1$. In the last case, there are three different situations. For example,

- $\displaystyle\sum_{n=-\infty\ n\neq 0}^{\infty} \frac{z^n}{n^2}$ converges for $|z| = 1$
- $\displaystyle\sum_{n=-\infty}^{\infty} z^n$ diverges for $|z| = 1$
- $\displaystyle\sum_{n=-\infty,\ n\neq 0}^{\infty} \frac{z^n}{n}$ converges everywhere on $|z| = 1$ except for $z = 1$.

We now show that every function analytic in an annulus has a Laurent series representation.

Theorem 9.1. (Laurent's Theorem) *Suppose $f(z)$ is analytic in the annulus $R_1 < |z-z_0| < R_2$. Then the representation*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$$

is valid throughout the annulus. Furthermore, the coefficients are given by

$$a_n = \frac{1}{2\pi i}\int_C \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\, d\zeta \quad (n = 0, \pm 1, \pm 2, \pm 3, \dots)$$

where C is any simple closed contour contained in the annulus that makes a complete counterclockwise revolution about the point z_0.

Proof. Let z be a point in the annulus. Construct two circles

$$C_1 : |z - z_0| = R_1' \ \text{ and } \ C_2 : |z - z_0| = R_2',$$

where R_1' and R_2' are such that $R_1 < R_1' < |z - z_0| < R_2' < R_2$. Then $f(z)$ is analytic in the closed annulus $R_1' \le |z - z_0| \le R_2'$ (see Figure 9.1). By

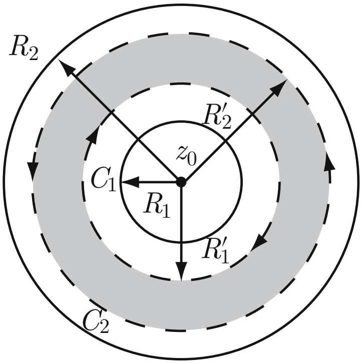


Figure 9.1.

Cauchy's integral formula for doubly connected domains,

$$f(z) = \frac{1}{2\pi i} \oint_{C_2} \frac{f(\zeta)}{\zeta - z}\, d\zeta - \frac{1}{2\pi i} \oint_{C_1} \frac{f(\zeta)}{\zeta - z}\, d\zeta. \tag{9.2}$$

Consider the first integral in (9.2). For ζ on C_2 and z in the annulus, we have $R_2' = |\zeta - z_0| > |z - z_0|$ and so, as in the proof of Theorem 8.8,

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

Consequently, we find

$$\frac{1}{2\pi i} \int_{C_2} \frac{f(\zeta)}{\zeta - z}\, d\zeta = \sum_{k=0}^{n-1} a_k (z - z_0)^k + R_n(z)$$

where $R_n(z) \to 0$ as $n \to \infty$ for $|z - z_0| < R_2'$, and a_k is given by

$$a_k = \frac{1}{2\pi i} \int_{C_2} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}}\, d\zeta \quad (k = 0, 1, 2, \dots).$$

Hence

$$\frac{1}{2\pi i} \int_{C_2} \frac{f(\zeta)}{\zeta - z}\, d\zeta = \sum_{k=0}^{\infty} a_k (z - z_0)^k, \quad |z - z_0| < R_2.$$

Observe that a_k is not in general equal to $f^{(k)}(z_0)/k!$ as was the case in the Taylor theorem because $f(z)$ may not be analytic at all points inside C_2.

Next consider the second integral in (9.2). For ζ on C_1 and z in the annulus $R_1' < |z - z_0| < R_2'$, we have $|\zeta - z_0| < |z - z_0|$ and so we seek an expression for $-1/(\zeta - z)$ in powers of $(\zeta - z_0)/(z - z_0)$. Accordingly, we write

$$\begin{aligned}
-\frac{1}{\zeta - z} &= \frac{1}{(z - z_0)[1 - (\zeta - z_0)/(z - z_0)]} \\
&= \sum_{k=0}^{n-1} \frac{1}{(\zeta - z_0)^{-k}} (z - z_0)^{-k-1} + \left(\frac{\zeta - z_0}{z - z_0}\right)^n \frac{1}{z - \zeta} \\
&= \sum_{k=-n}^{-1} \frac{1}{(\zeta - z_0)^{k+1}} (z - z_0)^{k} + \left(\frac{\zeta - z_0}{z - z_0}\right)^n \frac{1}{z - \zeta}.
\end{aligned}$$

Inserting this into the second integral in (9.2), we find that

$$-\frac{1}{2\pi i} \int_{C_1} \frac{f(\zeta)}{\zeta - z}\, d\zeta = \sum_{k=-n}^{-1} a_k (z - z_0)^k + S_n(z)$$

where

$$a_k = \frac{1}{2\pi i} \int_{C_1} \frac{f(\zeta)}{(\zeta - z)^{k+1}}\, d\zeta, \quad k = -1, -2, \ldots,$$

and

$$S_n(z) = \frac{1}{2\pi i} \int_{C_1} \frac{f(\zeta)}{z - \zeta} \left(\frac{\zeta - z_0}{z - z_0}\right)^n d\zeta.$$

Now for ζ on C_1,

$$|z - \zeta| = |z - z_0 - (\zeta - z_0)| \geq |z - z_0| - |\zeta - z_0| = |z - z_0| - R_1'.$$

Thus,

$$|S_n(z)| \leq \frac{1}{2\pi} \max_{\zeta \in C_1} |f(\zeta)| \frac{1}{|z - z_0| - R_1'} \left(\frac{R_1'}{|z - z_0|}\right)^n 2\pi R_1'.$$

Since $R_1' < |z - z_0|$, $S_n(z) \to 0$ as $n \to \infty$. Consequently,

$$-\frac{1}{2\pi i} \int_{C_1} \frac{f(\zeta)}{\zeta - z}\, d\zeta = \sum_{k=-\infty}^{-1} a_k (z - z_0)^k.$$

We have shown that for $R_1' < |z - z_0| < R_2'$,

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k + \sum_{k=-\infty}^{-1} a_k (z - z_0)^k, \tag{9.3}$$

where

$$a_k = \frac{1}{2\pi i} \int_{C_2} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}} \, d\zeta \quad (k = 0, 1, 2, \dots)$$

and

$$a_k = \frac{1}{2\pi i} \int_{C_1} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}} \, d\zeta \quad (k = -1, -2, \dots).$$

Now choose any simple closed contour contained in the annulus that makes a complete counterclockwise revolution about the point z_0. Since $f(\zeta)/(\zeta - z_0)$ is analytic in the region bounded by the closed contours C_1 and C (C_2 and C), Cauchy's integral formula implies that the coefficients may be computed by replacing C_1 and C_2 by C. Thus,

$$a_k = \frac{1}{2\pi i} \int_{C} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}} \, d\zeta \quad (k = 0, \pm 1, \pm 2, \pm 3, \dots). \tag{9.4}$$

Finally, since R_1' and R_2' may be chosen arbitrarily close to R_1 and R_2 respectively, the expression (9.3) (with a_k defined by (9.4)) is valid for all z in the annulus $R_1 < |z - z_0| < R_2$. ■

Remark 9.2. Observe that the series of positive powers of $z - z_0$ converges everywhere inside the circle $|z - z_0| = R_2$, whereas the series of negative powers of $z - z_0$ converges everywhere *outside* the circle $|z - z_0| = R_1$. The series of negative powers of $z - z_0$ is called the *principle part* of the Laurent expansion, while the series of positive powers is called the *analytic part.* ●

Remark 9.3. The Laurent expansion, like the power series expansion for analytic functions, is unique. Suppose that

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n = \sum_{n=-\infty}^{\infty} b_n (z - z_0)^n,$$

which is valid for $R_1 < |z - z_0| < R_2$. Then each series converges uniformly on a circle C contained in the annulus and enclosing z_0. Multiplying by $\frac{1}{2\pi i}(z - z_0)^k$ for any integer k, and integrating along C, we obtain

$$\sum_{n=-\infty}^{\infty} \frac{a_n}{2\pi i} \int_C (z - z_0)^{n+k} \, dz = \sum_{n=-\infty}^{\infty} \frac{b_n}{2\pi i} \int_C (z - z_0)^{n+k} \, dz. \tag{9.5}$$

Since

$$\frac{1}{2\pi i} \int_C (z - z_0)^m \, dz = \begin{cases} 1 & \text{if } m = -1 \\ 0 & \text{otherwise} \end{cases},$$

we get from (9.5)

$$a_{-k-1} = b_{-k-1} \quad \text{for each} \quad k \in \mathbb{Z}.$$

Hence, $a_k = b_k$ for every integer k showing that the Laurent expansion is unique. ●

Remark 9.4. The coefficients of a Laurent series are usually not found by the integral representation (9.4). In fact, determining the coefficients a_k by other means will enable us to evaluate the integral given in (9.4). Of particular interest is the coefficient a_{-1}, for this enables us to determine $\int_C f(\zeta)\,d\zeta$. We shall focus our attention on this part in later sections. ●

We now give some examples to show different methods for computing the coefficients of a Laurent series.

Example 9.5. To find the Laurent expansion for

$$f(z) = \frac{\sin z}{z^2} \quad (|z| > 0),$$

we expand $\sin z$ in a Maclaurin series. This observation leads to

$$f(z) = \frac{\sin z}{z^2} = \frac{1}{z^2}\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots\right) = \frac{1}{z} - \frac{z}{3!} + \frac{z^3}{5!} - \cdots \quad (|z| > 0).$$

Similarly, from the identity $e^u = 1 + u + u^2/2! + u^3/3! + \cdots$, we get

$$f(z) = e^{1/z} = 1 + \frac{1}{z} + \frac{1}{2!z^2} + \frac{1}{3!z^3} + \cdots \quad (|z| > 0).$$ ●

Example 9.6. Consider the function

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

which is analytic in $\mathbb{C}\setminus\{i,-i\}$. We first expand this function in a Laurent series valid in a deleted neighborhood of $z = i$. To do this, we consider

$$\frac{1}{z+i} = \frac{1}{2i+(z-i)} = \frac{1}{2i[1+(z-i)/2i]} = \frac{1}{2i}\sum_{n=0}^{\infty}(-1)^n\left(\frac{z-i}{2i}\right)^n$$

which is valid for $|z-i| < 2$. Hence, for $0 < |z-i| < 2$, we have

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

Similarly, to expand in a Laurent series valid in a deleted neighborhood of $z = -i$, we first write

$$\frac{1}{z-i} = \frac{1}{-2i+(z+i)} = -\frac{1}{2i[1-(z+i)/2i]} = -\frac{1}{2i}\sum_{n=0}^{\infty}\left(\frac{z+i}{2i}\right)^n,$$

which is valid for $|z+i| < 2$. Thus, for $0 < |z+i| < 2$,

$$f(z) = -\sum_{n=-1}^{\infty}\left(\frac{1}{2i}\right)^{n+2}(z+i)^n = -\sum_{n=-1}^{\infty}\left(-\frac{i}{2}\right)^{n+2}(z+i)^n. \qquad \bullet$$

Example 9.7. The function $f(z) = 1/[(z-1)(z-2)]$ is analytic in $\mathbb{C}\setminus\{1,2\}$. As in the previous example, we can expand this function in a Laurent series valid in a deleted neighborhood of $z=1$ or $z=2$. This observation leads to

$$f(z) = -\frac{1}{(z-1)[1-(z-1)]} = -\sum_{n=-1}^{\infty}(z-1)^n \quad (0 < |z-1| < 1)$$

and

$$f(z) = \frac{1}{(z-2)[1+(z-2)]} = \sum_{n=-1}^{\infty}(-1)^{n+1}(z-2)^n \quad (0 < |z-2| < 1).$$

But we can also expand $f(z)$ in Laurent series that are valid in different regions. For example, f has three Laurent series centered at 0:

(i) $|z| < 1$ (ii) $1 < |z| < 2$ (iii) $|z| > 2$.

(i) Suppose $|z| < 1$. Then (as $|z| < 1$ implies $|z| < 2$), we get

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

which is the Maclaurin series expansion for $f(z)$, i.e., no principal part.

(ii) Suppose $1 < |z| < 2$. Then

$$f(z) = -\frac{1}{2(1-z/2)} - \frac{1}{z(1-1/z)} = -\frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{z}{2}\right)^n - \frac{1}{z}\sum_{n=0}^{\infty}\left(\frac{1}{z}\right)^n.$$

Here we are using the fact that $|z/2| < 1$ and $|1/z| < 1$. Hence

$$f(z) = -\sum_{n=-\infty}^{-1} z^n - \sum_{n=0}^{\infty}\frac{1}{2^{n+1}}z^n \quad \text{for } 1 < |z| < 2.$$

(iii) Suppose $|z| > 2$. Then (as $|z| > 2$ implies $|2/z| < 1$ and $|1/z| < 1$)

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

which has no analytic part. $\bullet$

Questions 9.8.

1. What similarities are there between Laurent series and power series?
2. If the Laurent series for $f(z)$ converges on the boundary of an annulus, is $f(z)$ analytic on the boundary?
3. Where does a Laurent series converge uniformly?
4. Can a function be analytic only in a rectangular strip?
5. Can $f(z)$ and $f(1/z)$ be analytic at the same set of points?
6. What can be said about the sum of two Laurent series? The product?
7. If $f(z) = \sum_{n=-\infty}^{\infty} a_n(z-z_0)^n$, under what circumstances does
$$f'(z) = \sum_{n=-\infty}^{\infty} na_n(z-z_0)^{n-1}?$$
8. Does the function $f(z) = \overline{z}$ have Laurent series valid in some domain $0 < |z| < \delta$? How about if $f(z) = \operatorname{Re} z$ or $\operatorname{Im} z$ or $|z|^2$?

Exercises 9.9.

1. Expand $f(z) = \dfrac{1}{(z+1)(z^2+2)}$ in Laurent series valid for
   (i) $1 < |z| < \sqrt{2}$ (ii) $|z| > \sqrt{2}$ (iii) $|z| < 1$.
2. Expand $f(z) = \dfrac{3z-1}{z^2-2z-3}$ in Laurent series valid for
   (i) $1 < |z| < 3$ (ii) $|z| > 3$ (iii) $|z| < 1$.
3. Find the Laurent series for $f(z) = \dfrac{z-12}{z^2+z-6}$ valid for
   (i) $1 < |z-1| < 4$ (ii) $|z-1| > 1$ (iii) $|z-1| < 4$.
4. Expand $f(z) = e^{z^2} + e^{1/z^2}$ in a Laurent series valid for $|z| > 0$.
5. For $0 < |a| < |b|$, expand $f(z) = \dfrac{1}{(z-a)(z-b)}$ in Laurent series valid for
   (i) $|z| < |a|$ (ii) $|a| < |z| < |b|$
   (iii) $|z| > |b|$ (iv) $0 < |z-a| < |a-b|$
   (v) $0 < |z-b| < |a-b|$.
6. If $0 < |a| < |b|$, expand
$$f(z) = \frac{1}{z(z-a)(z-b)}$$
   as a Laurent series valid in
   (i) $0 < |z| < |a|$ (ii) $|a| < |z| < |b|$ (iii) $|z| > |b|$.

7. Expand
$$f(z) = \frac{z^2 + 9z + 11}{(z+1)(z+4)}$$
as a Laurent series about $z = 0$ valid when

(i) $|z| < 1$ (ii) $1 < |z| < 4$ (iii) $|z| > 4$.

8. Find the principal part for the following Laurent series.
   (a) $\dfrac{z^2}{z^4 - 1}$ $(0 < |z - i| < \sqrt{2})$
   (b) $\dfrac{z^2}{z^4 - 1}$ $(0 < |z + i| < \sqrt{2})$
   (c) $\dfrac{e^z}{z^4}$ $(|z| > 0)$
   (d) $\dfrac{\sin z}{z^4}$ $(|z| > 0)$
   (e) $\dfrac{1}{\tan^2 z} - \dfrac{1}{z^2}$ $(0 < |z| < \pi/2)$.
9. Expand the following in a Laurent series valid in the region indicated.
   (a) $z^n e^{1/z}$ $(|z| > 0)$ (b) $e^{1/(z-1)}$ $(|z| > 1)$.
10. Express $\sin z \sin(1/z)$ in a Laurent series valid for $|z| > 0$.
11. Use series division to find the principal part in a neighborhood of the origin for the function $e^z/(1 - \cos z)^2$.

9.2 Classification of Singularities

A single-valued function is said to have a *singularity* at a point if the function is not analytic at the point while every neighborhood of that point contains at least one point at which the function is analytic. For instance, $f(z) = \overline{z}$ is nowhere analytic. Note that we do not say that every point of $\mathbb{C}$ is a singularity for $f(z)$. Basically, there are two types of singularities

(i) isolated singularity
(ii) non-isolated singularity.

If the function is analytic in some deleted neighborhood of the point, then the singularity is said to be isolated. For example, consider

(i) $f_1(z) = \begin{cases} z+1 & \text{for } z \neq 0 \\ 3 & \text{for } z = 0 \end{cases}$
(ii) $f_2(z) = 1/z$ for $z \neq 0$
(iii) $f_3(z) = \sin(1/z)$ for $z \neq 0$.

For each of the functions defined above, the point $z = 0$ is an isolated singularity.

Recall from Section 4.3 that $f(z) = \operatorname{Log} z$, the principal logarithm, is analytic for $z \in \Omega = \mathbb{C} \setminus \{x + iy : y = 0, x \leq 0\}$. The point $z = 0$ is a branch point of $\operatorname{Log} z$ since every deleted neighborhood of $z = 0$ contains points on the negative real axis. *We say that a singularity at $z = z_0$ of f is non-isolated if every neighborhood of a contains at least one singularity of f other than a.* For example, $z = 0$ as well as every point on the negative real axis is a non-isolated singularity of the principal logarithm given by

$$\operatorname{Log} z = \ln |z| + i\operatorname{Arg} z, \quad -\pi < \operatorname{Arg} z < \pi,$$

since every neighborhood of $z = x$ $(x \leq 0)$ contains points on the negative real axis on which $\operatorname{Log} z$ is not analytic. The behavior of a function $f(z)$ near an isolated singularity z_0 can be described by considering the limiting value of $\lim_{z \to z_0} f(z)$. Then there are three possibilities:

(i) *$f(z)$ may be bounded in a deleted neighborhood of z_0*. For instance, in the above example, $f_1(z)$ is bounded in a deleted neighborhood of the origin. This is an uninteresting example because $f_1(z)$ can be made analytic by defining $f_1(0) = 1$. Another example of this type may be given by

$$g_1(z) = \frac{\sin z}{z} \quad \text{for } z \neq 0.$$

(ii) *$f(z)$ may approach ∞ as z approaches z_0*. For instance,

$$f_2(z) = \frac{1}{z} \to \infty \quad \text{as} \quad z \to 0.$$

Another example of this type may be given by $g_1(z) = \frac{1}{z^n}$ $(z \neq 0)$, where $n \in \mathbb{N}$ is fixed.

(iii) *$f(z)$ may satisfy neither* (i) *nor* (ii). For instance, consider

$$f(z) = e^{1/z}, \quad z \neq 0.$$

For $z = x$, $x \neq 0$ real, note that

$$\lim_{x \to 0^+} e^{1/x} = \infty \text{ and } \lim_{x \to 0^-} e^{1/x} = 0.$$

Moreover, for $z = iy$ $(y \in \mathbb{R} \setminus \{0\})$, we note that

$$f(iy) = e^{1/(iy)} = e^{-i/y}$$

which lies always on the unit circle. Consequently, $\lim_{z \to 0} e^{1/z}$ does not exist. Thus, in a neighborhood of the origin, $e^{1/z}$ is neither bounded nor does it approach ∞. A similar argument continues to hold for $\sin(1/z)$ in the neighborhood of the origin. Indeed, we consider two sequences:

$$z_n = i/n \to 0 \ \text{ and } \ z_n' = 1/n \to 0 \ as\ n \to \infty.$$

Then

$$\sin(1/z_n) = \sin(-in) = \frac{e^n - e^{-n}}{2i} \to \infty,$$

whereas $\{\sin(1/z_n')\} = \{\sin n\}$ is clearly a bounded sequence. It follows that $z = 0$ is an essential singularity of $\sin(1/z)$.

It follows in general that if a function has one non-isolated singularity, then it will have many singularities, although not necessarily non-isolated. This is demonstrated by the following example. The function

$$f(z) = \frac{1}{\sin(1/z)}$$

has a singularity at the origin because $\lim_{z\to 0} f(z)$ does not exist either as a finite limit or as an infinite limit. Note that $\lim_{z\to 0}\sin(1/z)$ does not exist because

$$z_n = \frac{1}{n\pi} \to 0 \ \text{ and } \ z_n' = \frac{1}{n\pi + \pi/2} \to 0,$$

whereas $\sin(z_n) = 0$ and $\sin(z_n') = \cos(n\pi) = (-1)^n$. Furthermore, the zeros of $\sin(1/z)$ are given by $z_n = 1/(n\pi)$, $n \in \mathbb{Z}\backslash\{0\}$. These points are also singularities of $f(z)$. Notice that $|z_n| \to 0$ as $n \to \infty$ so that each deleted neighborhood of the origin contains a singularity of $f(z)$. Consequently, the singularity at $z = 0$ is a non-isolated singularity of $f(z)$.

Observe that the "clever" way we phrased (iii) makes every isolated singularity satisfy either (i), (ii) or (iii).

Each of the three kinds of isolated singularities have important characterizations which will be obtained in the next three theorems.

Theorem 9.10. (Riemann's Theorem) *Suppose that $f(z)$ has an isolated singularity at $z = z_0$ and is bounded in some deleted neighborhood of z_0. Then $f(z)$ can be defined at z_0 in such a way as to be analytic at z_0.*

Proof. Assume the hypothesis. Then, for some $R > 0$, $f(z)$ is analytic in the punctured disk

$$0 < |z - z_0| \le R.$$

Given a point z_1 inside this disk, choose $r > 0$ so that $r < |z_1 - z_0| < R$. The function $f(z)$ is analytic in the annulus bounded by the two circles $C : |z - z_0| = R$ and $C_1 : |z - z_0| = r$ (see Figure 9.2). Hence by Cauchy's integral formula,

$$f(z_1) = \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta - z_1}\, d\zeta - \frac{1}{2\pi i}\oint_{C_1} \frac{f(\zeta)}{\zeta - z_1}\, d\zeta. \tag{9.6}$$

Observe that the value of (9.6) is independent of the choice of r. We will prove that the last integral on the right of (9.6) is zero by showing that its absolute

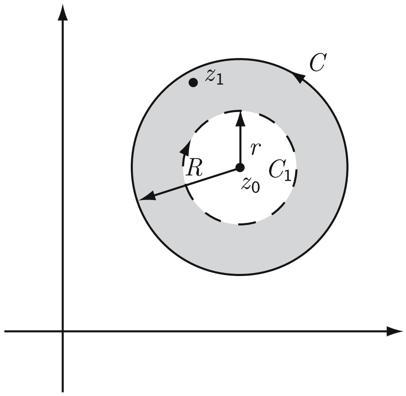


Figure 9.2.

value can be made arbitrarily small for sufficiently small values of r. Since $f(z)$ is bounded in the disk, say $|f(z)| \leq M$, we have

$$\left|\frac{1}{2\pi i}\int_{C_1}\frac{f(\zeta)}{\zeta - z_1}\,d\zeta\right| \leq \frac{M}{2\pi}\int_{C_1}\frac{|d\zeta|}{|\zeta - z_1|} \leq \frac{M}{2\pi}\,\frac{2\pi r}{|z_1 - z_0| - r} \to 0 \text{ as } r \to 0,$$

and so (9.6) may be written as

$$f(z_1) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - z_1}\,d\zeta. \tag{9.7}$$

Since z_1 is arbitrary in the disk $0 < |z - z_0| < R$, we may write (9.7) as

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - z}\,d\zeta, \tag{9.8}$$

valid for all z, $0 < |z - z_0| < R$. In the proof of Theorem 8.3 it was shown that the integral on the right side of (9.8) represents an analytic function for $|z - z_0| < R$. Hence by defining

$$f(z_0) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - z_0}\,d\zeta,$$

the function $f(z)$ becomes analytic in the whole disk $|z - z_0| < R$. ∎

Corollary 9.11. *If $f(z)$ has an isolated singularity at $z = z_0$ and is bounded in some neighborhood of z_0, then $\lim_{z\to z_0} f(z)$ exists.*

Proof. Since an analytic function is continuous, $f(z_0)$ had to have been defined in Theorem 9.10 in such a way that $\lim_{z\to z_0} f(z) = f(z_0)$. In particular, $\lim_{z\to z_0} f(z)$ exists. ∎

We have shown that functions satisfying the conditions of Theorem 9.10 have a singularity simply because they have not been defined "properly" at the point in question. *If $f(z)$ has an isolated singularity at z_0, the singularity is said to be a removable singularity if* $\lim_{z\to z_0} f(z)$ *exists.* Theorem 9.10 says that a function which is analytic and bounded in a deleted neighborhood of a point has at worst a removable singularity at the point. For all practical purposes, we may consider such a function to be analytic. Thus when we speak of $(\sin z)/z$ as an entire function, it will be understood that $f(0) = \lim_{z\to 0}(\sin z)/z = 1$. Note that the Maclaurin expansion for this function is

$$f(z) = \frac{\sin z}{z} = \frac{1}{z}\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots\right) = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} + \cdots, \quad |z| > 0.$$

A function $f(z)$, analytic in a deleted neighborhood of $z = z_0$, has a pole of order k (k a positive integer) if

$$\lim_{z\to z_0} (z - z_0)^k f(z) = A \neq 0, \infty.$$

We now characterize singularities of the form (ii) quoted in the beginning.

Theorem 9.12. *If $f(z)$ has an isolated singularity at $z = z_0$ and $f(z) \to \infty$ as $z \to z_0$, then $f(z)$ has a pole at $z = z_0$.*

Proof. Suppose that $f(z) \to \infty$ as $z \to z_0$. Then for a given $R > 0$ there exists a $\delta > 0$ such that $f(z)$ is analytic for $0 < |z - z_0| < \delta$ and

$$|f(z)| > R \quad \text{whenever} \quad 0 < |z - z_0| < \delta.$$

In particular, $f(z) \neq 0$ for $0 < |z-z_0| < \delta$ and so, $g(z) = 1/f(z)$ is analytic and bounded by $1/R$ in this deleted neighborhood of z_0. By Theorem 9.10, $g(z)$ has a removable singularity at z_0, and we may write $g(z)$ as (using $\lim_{z\to z_0} g(z) = 0$)

$$g(z) = \frac{1}{f(z)} = a_1(z - z_0) + a_2(z - z_0)^2 + \cdots, \quad 0 < |z - z_0| < \delta.$$

Moreover, $g(z) \neq 0$ for $0 < |z - z_0| < \delta$, and so not all the coefficients of $g(z)$ are zero. This means that there is a $k \geq 1$ such that a_k is the first nonzero coefficient of $g(z)$. Then

$$g(z) = \frac{1}{f(z)} = a_k(z - z_0)^k + a_{k+1}(z - z_0)^{k+1} + \cdots,$$

so that

$$\frac{1}{(z - z_0)^k f(z)} = a_k + a_{k+1}(z - z_0) + \cdots \to a_k \quad \text{as } z \to z_0, \tag{9.9}$$

and therefore,

$$\lim_{z\to z_0} (z - z_0)^k f(z) = \frac{1}{a_k}.$$

Hence, by the definition, $f(z)$ has a pole of order k at $z = z_0$. ■

Remark 9.13. If $f(z)$ has a pole at z_0, it follows from the definition that there exists a $k \geq 1$ such that $(z-z_0)^k f(z) \to A \neq 0$ as $z \to z_0$. Thus, for $z \to z_0$,

$$\left|\frac{1}{f(z)}\right| = \left|\frac{(z-z_0)^k}{(z-z_0)^k f(z)}\right| \to \left|\frac{0}{A}\right| = 0 \quad \text{as } z \to z_0$$

and hence, $f(z) \to \infty$ as $z \to z_0$. Thus, Theorem 9.12 gives a necessary and sufficient condition for an isolated singularity to be a pole. ●

Corollary 9.14. *If $f(z)$ has a pole at $z = z_0$, then $f(z)$ may be expressed as*

$$f(z) = \sum_{n=-k}^{\infty} b_n (z-z_0)^n,$$

where k is the order of the pole.

Proof. We use the notation of Theorem 9.12. By (9.9), the function

$$\frac{1}{f(z)(z-z_0)^k} = a_k + a_{k+1}(z-z_0) + \cdots \qquad (a_k \neq 0)$$

is analytic at $z = z_0$. Since $a_k \neq 0$, a continuity argument shows that there is a neighborhood of z_0 in which

$$\frac{1}{f(z)(z-z_0)^k} \neq 0.$$

Hence, $f(z)(z-z_0)^k$ is analytic at z_0 and the expansion

$$f(z)(z-z_0)^k = \sum_{m=0}^{\infty} c_m (z-z_0)^m \qquad \left(c_0 = \frac{1}{a_k} \neq 0\right)$$

is valid in some neighborhood of z_0. That is,

$$f(z) = \sum_{m=0}^{\infty} c_m (z-z_0)^{m-k}. \tag{9.10}$$

Upon setting $n = m-k$ and $b_n = c_{m+k}$ in (9.10), we get the desired form. ■

An isolated singularity that is neither removable nor a pole is said to be an (isolated) essential singularity. Equivalently, a function $f(z)$ which is analytic in a deleted neighborhood of $z = z_0$ has an essential singularity at $z = z_0$ if there exists no nonnegative integer k for which $\lim_{z\to z_0}(z-z_0)^k f(z)$ exists (either as a finite value or as an infinite value). The behavior of a function in the neighborhood of an isolated essential singularity is most surprising.

Theorem 9.15. (Casorati–Weierstrass) *If $f(z)$ has an essential singularity at $z = z_0$, then $f(z)$ comes arbitrarily close to every complex value in each deleted neighborhood of z_0.*

Proof. Let f have an essential singularity at z_0. Then f is analytic throughout a deleted neighborhood of z_0. Suppose, for some complex number a, that

$$|f(z) - a| \geq \epsilon > 0$$

for all z in a punctured disk $0 < |z - z_0| < \delta$. Set $g(z) = 1/(f(z) - a)$. Then

$$|g(z)| = \left|\frac{1}{f(z) - a}\right| \leq \frac{1}{\epsilon} \quad \text{for } 0 < |z - z_0| < \delta.$$

Thus, by Theorem 9.10, $g(z)$ has a removable singularity at $z = z_0$ and we may write

$$g(z) = \frac{1}{f(z) - a} = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots .$$

Observe that $\lim_{z \to z_0} 1/(f(z) - a) = a_0$. There are two cases to consider.

Case 1) $a_0 \neq 0$: Then $\lim_{z \to z_0} f(z) = 1/a_0 + a$, and $f(z)$ has a removable singularity at $z = z_0$.

Case 2) $a_0 = 0$: Suppose a_k is the first nonzero coefficient. Then

$$\frac{1}{(f(z) - a)(z - z_0)^k} = a_k + a_{k+1}(z - z_0) + \cdots .$$

In this case,

$$\lim_{z \to z_0} (z - z_0)^k (f(z) - a) = \lim_{z \to z_0} (z - z_0)^k f(z) = \frac{1}{a_k},$$

and $f(z)$ has a pole of order k at $z = z_0$. Since our hypothesis excludes both cases, the inequality $|f(z) - a| \geq \epsilon > 0$ cannot be true and the desired conclusion follows. ∎

Corollary 9.16. *Suppose $f(z)$ has an essential singularity at $z = z_0$. Given any complex number a, there exists a sequence $\{z_n\}$ such that $z_n \to z_0$ and $f(z_n) \to a$.*

Proof. Choose a sequence $\{\delta_n\}$ for which $\delta_n > 0$ for each value of n and $\lim_{n \to \infty} \delta_n = 0$. By Theorem 9.15, we can find a sequence of points $\{z_n\}$ such that $|f(z_n) - a| < 1/n$ for $0 < |z_n - z_0| < \delta_n$. Thus, $f(z_n) \to a$ as $z_n \to z_0$. ∎

A generalization of Theorem 9.15 is described by Picard in the following form. For details of this result, we refer to advanced texts, see Hille [Hi].

Theorem 9.17. (Picard's Great Theorem) *In the neighborhood of an isolated essential singularity, a function assumes each complex value, with one possible exception, infinitely often.*

The function $e^{1/z}$ has an isolated essential singularity at $z = 0$, and never assumes the value 0. We wish to show that $e^{1/z}$ assumes every other value in a neighborhood of the origin infinitely often. The behavior near an essential singularity is strange. Indeed, if $c \neq 0$, then the solutions of

$$e^{1/z} = c$$

are given by

$$z_n = \frac{1}{\log c} = \frac{1}{\operatorname{Log} c + 2n\pi i}, \quad n \in \mathbb{Z}.$$

Observe that infinitely many z_n are contained in every neighborhood of the origin.

Isolated singularities at $z = \infty$. We may also refer to the point at ∞ as being an isolated singularity. A function $f(z)$ has an isolated singularity at $z = \infty$ if $f(z)$ is analytic in a deleted neighborhood of ∞, that is there exists a real number R such that $f(z)$ is analytic for $R < |z| < \infty$. Note that $f(z)$ is analytic for $|z| > R$ if and only if $f(1/z)$ is analytic for $|1/z| < R$. Hence, $f(z)$ has an isolated singularity at $z = \infty$ if and only if $f(1/z)$ has an isolated singularity at $z = 0$. Moreover, we make the definition that the singularity of $f(z)$ at $z = \infty$ is removable, a pole, or essential according as the singularity of $f(1/z)$ at $z = 0$ is removable, a pole, or essential. For example, the function

1. $f(z) = z^2 + 1$ has a pole of order 2 at $z = \infty$ because $f(1/z) = (1/z^2) + 1$ has a pole of order 2 at $z = 0$.
2. $f(z) = e^z$ has an isolated essential singularity at $z = \infty$ because $f(1/z) = e^{1/z}$ has an isolated essential singularity at $z = 0$.
3. $f(z) = 1/[z(z^2 + 4)]$ has simple poles at $z = 0, \pm 2i$. Now

$$f(1/z) = \frac{z^3}{1 + 4z^2},$$

showing that $z = 0$ is a point of analyticity for $f(1/z)$. In fact, it has zero of order 3 at $z = 0$.

With this definition, we can examine the behavior of entire functions. If $f(z) = \sum_{n=0}^{k} a_n z^n$ is a polynomial of degree k, then

$$g(z) = f\left(\frac{1}{z}\right) = \frac{a_k}{z^k} + \frac{a_{k-1}}{z^{k-1}} + \cdots + \frac{a_1}{z} + a_0 \quad \text{and} \quad \lim_{z \to 0} z^k f\left(\frac{1}{z}\right) = a_k \neq 0$$

so that $g(z)$ has a pole of order k at $z = 0$. Hence, $f(z)$ has a pole of order k at $z = \infty$. Theorem 8.38 is now seen to be a special case of Theorem 9.15. That is, the polynomial $f(z) \to \infty$ as $z \to \infty$ because $f(z)$ has a pole at $z = \infty$. Note that the entire function $f(z)$ has a pole of order k at $z = \infty$ if and only if

$$\lim_{z \to 0} z^k f\left(\frac{1}{z}\right) = \lim_{z \to \infty} \frac{f(z)}{z^k} = A \neq 0, \infty. \tag{9.11}$$

Moreover, if (9.11) is satisfied, then (by Theorem 8.35) $f(z)$ is a polynomial of degree k. We summarize the discussion as

Theorem 9.18. *Let $f(z)$ be a nonconstant entire function. Then $f(z)$ is a polynomial if and only if $f(z) \to \infty$ as $z \to \infty$.*

Thus a transcendental entire function cannot have a pole at $z = \infty$. That is, transcendental entire functions must have essential singularities at $z = \infty$.

An interesting comparison can now be made between Theorem 8.31 and Theorem 9.15. Theorem 8.31 merely asserts that an entire function comes arbitrarily close to every complex value. Theorem 9.15 says that a transcendental entire function comes arbitrarily close to every complex value outside of every circle $|z| = R$. Of course the behavior of a non-transcendental entire function (a polynomial) has already been fully discussed.

In the Laurent series expansion (see Theorem 9.1), if $R_1 = 0$, then the point z_0 becomes an isolated singularity of $f(z)$. In view of this, the Laurent series allows us to classify the type of isolated singularity at z_0. Suppose that f has an isolated singularity at z_0. Then $f(z)$ is analytic throughout a deleted neighborhood of z_0, that is $f(z)$ possesses a Laurent series expansion

$$f(z) = \sum_{n=-\infty}^{-1} a_n (z - z_0)^n + \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

valid for $0 < |z - z_0| < \delta$ for some $\delta > 0$. Then the following situations arise.

(i) **No principal part**: In this case $a_n = 0$ for all $n < 0$. So the above Laurent series simply reduces to

$$f(z) = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots, \quad 0 < |z - z_0| < \delta.$$

It follows that $\lim_{z \to z_0} f(z) = a_0$, i.e., f is bounded in a deleted neighborhood of z_0. In other words, f has a removable singularity at z_0 if and only if $a_n = 0$ for all $n < 0$.

(ii) **The principal part consists of a finite number of terms**: In this case, $a_n = 0$ for all $n < -m$ and $a_{-m} \neq 0$ for some $m \geq 1$. So the Laurent series about z_0 reduces to

$$f(z) = \frac{a_{-m}}{(z - z_0)^m} + \cdots + \frac{a_{-1}}{z - z_0} + \sum_{n=0}^{\infty} a_n (z - z_0)^n, \tag{9.12}$$

which is valid for $0 < |z - z_0| < \delta$ so that

$$\lim_{z \to z_0} (z - z_0)^m f(z) = a_{-m} \neq 0.$$

If we rewrite (9.12) in the form

$$\begin{aligned} f(z) &= \frac{a_{-m}}{(z-z_0)^m}\left[1 + \frac{a_{-m+1}}{a_{-m}}(z-z_0) + \cdots \right. \\ &\qquad \left. + \frac{a_{-1}}{a_{-m}}(z-z_0)^{m-1} + \sum_{n=0}^{\infty} \frac{a_n}{a_{-m}}(z-z_0)^{n+m}\right] \\ &= \frac{a_{-m}}{(z-z_0)^m}\left[1 + g(z)\right], \end{aligned}$$

we see that $g(z) \to 0$ as $z \to z_0$. Thus, given $\epsilon = 1/2$, there exists a $\delta > 0$ such that

$$|g(z)| < 1/2 \quad \text{whenever } |z - z_0| < \delta$$

and therefore, $|1 + g(z)| \geq 1 - |g(z)| \geq 1/2$ for $|z - z_0| < \delta$. This observation shows that

$$|f(z)| \geq \frac{|a_{-m}|}{|z-z_0|^m}|1 + g(z)| > \frac{|a_{-m}|}{2|z-z_0|^m} > \frac{|a_{-m}|}{2\delta^m}$$

which, for $\delta \to 0$, implies that $f(z) \to \infty$ as $z \to z_0$. In particular, $f(z)$ behaves like $a_{-m}/(z-z_0)^m$ for z near z_0. In other words, *$f(z)$ has a pole of order m at z_0 if and only if there exists an $m \in \mathbb{N}$ such that $a_{-m} \neq 0$ and $a_n = 0$ for all n less than $-m$.*

(iii) **The principal part consists of infinitely many terms**: In this case, $a_n \neq 0$ for infinitely many negative integer values of n. Thus by a process of elimination, *the principal part has infinitely many terms if and only if the singularity at z_0 is an essential singularity.*

From the above discussion, when expanding in a Laurent series about an isolated singularity, we are sometimes interested only in the principal part. If $f(z)$ has a simple pole at $z = z_0$, then the principal part is particularly easy to determine. For then $\lim_{z \to z_0}(z - z_0)f(z)$ must exist, and we may set

$$(z - z_0)f(z) = a_{-1} + a_0(z - z_0) + a_1(z - z_0)^2 + \cdots \qquad (a_{-1} \neq 0).$$

The principal part is then seen to be $a_{-1}/(z - z_0)$.

Examples 9.19. (i) For instance, consider $f(z) = z/(z^2+4)$. Suppose that we wish to find the principal part of the Laurent expansion valid in a deleted neighborhood of $z = 2i$. Since

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

$f(z)$ has a simple pole at $z = 2i$. Therefore,

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

where $g(z)$ is analytic at $z = 2i$.

(ii) If $f(z) = 1/(z^4 + 1)$, then the solution set of $z^4 + 1 = 0$ is given by

$$z_k = (-1)^{1/4} = e^{i(\pi+2k\pi)/4}, \quad k = 0, 1, 2, 3,$$

and these are the simple poles of $f(z)$. For example, as in the previous example, we get

$$\lim_{z\to z_k} \frac{z - z_k}{z^4 + 1} = \lim_{z\to z_k} \frac{1}{4z^3} = \frac{1}{4z_k^3} = \frac{z_k}{4z_k^4} = -\frac{z_k}{4}.$$

Then,

$$f(z) = \frac{-z_k/4}{z - z_k} + g_k(z) \quad \text{for} \quad k = 0, 1, 2, 3,$$

where each $g_k(z)$ is analytic at $z = z_k$. In particular, at $z_0 = e^{i\pi/4}$

$$f(z) = \frac{-(1+i)/\sqrt{2}}{4(z - e^{i\pi/4})} + g_0(z),$$

where $g_0(z)$ is analytic at $z_0 = e^{i\pi/4} = (1+i)/\sqrt{2}$.

(iii) Let us now find all singularities for the function $f(z) = \cot \pi z$, and determine the principal part of the Laurent expansion about each singularity. Since $\sin \pi z$ has simple zeros at $z = n$ ($n \in \mathbb{Z}$), the function $f(z) = \cos \pi z / \sin \pi z$ has simple poles at $z = n$. In a deleted neighborhood of $z = n$, we have

$$f(z) = \frac{a_{-1}}{z - n} + g_n(z), \quad 0 < |z - z_0| < 1,$$

where $g_n(z)$ is analytic at $z = n$. It remains to determine a_{-1}. From the identity

$$\sin \pi z = (-1)^n \sin \pi(z - n),$$

we get (as f has simple pole at each $z = n$)

$$a_{-1} = \lim_{z\to n} (z - n) f(z) = \frac{1}{\pi} \lim_{z\to n} \frac{\pi(z - n)}{\sin \pi(z - n)} \frac{\cos \pi z}{(-1)^n} = \frac{1}{\pi}$$

and the principal part of $f(z)$ at each $z = n$ is

$$\frac{1/\pi}{z - n}.$$

(iv) If $f(z) = \dfrac{2 + e^z}{\sin z + z \cos z}$, then $z = 0$ is an isolated singularity and

$$\lim_{z\to 0} z f(z) = \lim_{z\to 0} \frac{2 + e^z}{z^{-1} \sin z + \cos z} = \frac{3}{2} \neq 0,$$

which shows that $z = 0$ is a simple pole for $f(z)$. Similarly, we see that

$$g(z) = \frac{1 + \sin z}{\cos z - 1 + \sin z}$$

has a simple pole at $z = 0$, because

$$\lim_{z \to 0} z g(z) = \lim_{z \to 0} \frac{1 + \sin z}{\dfrac{\cos z - 1}{z} + \dfrac{\sin z}{z}} = 1.$$

(v) Let us discuss the nature of the singularity of $f(z) = (z+1)^{-4} \sin \pi z$ at $z = -1$ and write down the principal part of it. To do this, we first observe that $f(z)$ has a pole of order 3 at $z = -1$, and f is analytic for $0 < |z + 1| < \infty$. It follows that

$$f(z) = (z+1)^{-4} \sum_{n=0}^{\infty} a_n (z+1)^n,$$

where $a_n = g^{(n)}(-1)/n!$, with $g(z) = \sin \pi z$. Note that

$$\begin{aligned} g'(z) &= \pi \cos \pi z = \pi \sin(\pi z + \pi/2), \\ g''(z) &= \pi^2 \cos(\pi z + \pi/2) = \pi^2 \sin(\pi z + 2(\pi/2)), \\ \vdots\ &= \qquad \vdots \\ g^{(n)}(z) &= \pi^n \sin(\pi z + n\pi/2), \end{aligned}$$

and so

$$g^{(n)}(-1) = \pi^n \sin(-\pi + n\pi/2) = \begin{cases} 0 & \text{if } n = 2k \\ -\pi^n (-1)^k & \text{if } n = 2k+1,\ k = 0, 1, \dots. \end{cases}$$

Thus,

$$f(z) = (z+1)^{-4} \sum_{k=0}^{\infty} \frac{(-1)^{k+1} \pi^{2k+1}}{(2k+1)!} (z+1)^{2k+1}$$

so that $f(z)$ has a pole of order 3 at $z = -1$. From this expansion, one can easily write down the principal part. ●

Examples 9.20. We wish to characterize all rational functions which have a removable singularity at ∞.

To do this, we let $f(z) = p(z)/q(z)$, where p and q are polynomials. Then

$f(z)$ has a removable singularity at ∞

$$\begin{aligned} &\iff f(1/z) \text{ has a removable singularity at } z = 0 \\ &\iff |f(1/z)| \le M \text{ for } 0 < |1/z| < \epsilon, \text{ for some } \epsilon > 0 \text{ and } M > 0 \\ &\iff |f(z)| \le M \text{ for } |z| > 1/\epsilon \\ &\iff \deg p(z) \le \deg q(z). \end{aligned}$$

Suppose that we wish to characterize those rational functions which have a pole of order k at ∞. To do this, we proceed as follows: set $f(z) = p(z)/q(z)$, where p and q are polynomials of degree m and n, respectively. Then, with $p(z) = \sum_{j=0}^{m} p_j z^j$ and $q(z) = \sum_{j=0}^{n} q_j z^j$,

$f(z)$ has a pole of order k at ∞

$$
\begin{aligned}
&\iff f(1/z) \text{ has a pole of order } k \text{ at } z = 0 \\
&\iff f(1/z) = \frac{p(1/z)}{q(1/z)} = \sum_{j=-k}^{\infty} a_j z^j \\
&\iff \sum_{j=0}^{m} p_j \frac{1}{z^j} = \left(\sum_{j=0}^{n} q_j \frac{1}{z^j}\right)\left(\sum_{j=-k}^{\infty} a_j z^j\right) \\
&\iff m = n + k, \text{ i.e., } k = \deg p(z) - \deg q(z).
\end{aligned}
$$

●

Division by power series furnishes us with another method for determining the principal part. Suppose

$$f(z) = \frac{a_0 + a_1 z + \cdots}{b_0 + b_1 z + \cdots} \quad (b_0 \neq 0).$$

Then $f(z)$ is analytic at $z = 0$ and may be expanded in a series of the form

$$\frac{a_0 + a_1 z + \cdots}{b_0 + b_1 z + \cdots} = c_0 + c_1 z + \cdots,$$

which is valid in a neighborhood of origin. Using series multiplication, previously discussed, we have

$$a_0 + a_1 z + \cdots = (b_0 + b_1 z + \cdots)(c_0 + c_1 z + \cdots) = b_0 c_0 + (b_1 c_0 + c_1 b_0) z + \cdots.$$

We can now compute the c_k recursively by the equations

$$
\begin{aligned}
a_0 &= b_0 c_0, \\
a_1 &= b_0 c_1 + b_1 c_0, \\
&\vdots \\
a_n &= b_0 c_n + b_1 c_{n-1} + \cdots + b_n c_0.
\end{aligned}
$$

This method may be viewed as "long division". That is,

$$
\begin{array}{r|l}
 & \dfrac{a_0}{b_0} + \cdots \\
\hline
(b_0 + b_1 z + \cdots & a_0 + a_1 z + \cdots \\
 & a_0 + \dfrac{a_0 b_1}{b_0} z + \cdots \\
\cline{2-2}
\end{array}
.
$$

We shall use this method to find the principal part of

$$f(z) = \frac{\pi \cot \pi z}{z^4}$$

valid in a deleted neighborhood of the origin. We have

$$\begin{aligned} f(z) &= \frac{\pi \cos \pi z}{z^4 \sin \pi z} = \frac{1}{z^5}\left(\frac{1-(\pi z)^2/2!+(\pi z)^4/4!-\cdots}{1-(\pi z)^2/3!+(\pi z)^4/5!-\cdots}\right) \\ &= \frac{1}{z^5}\left(1+\frac{\pi^2}{3}z^2-\frac{\pi^4}{45}z^4+\cdots\right). \end{aligned}$$

Thus the principal part of $f(z)$ is

$$\frac{1}{z^5}+\frac{\pi^2/3}{z^3}-\frac{\pi^4/45}{z}.$$

Questions 9.21.

1. Can a function have infinitely many isolated singularities in the plane? In a bounded region? In a compact set?
2. Given a function $f(z)$, does there exist a real number M such that no pole of $f(z)$ has order greater than M?
3. Can a function have poles at a preassigned sequence of points?
4. Can a function have essential singularities at the preassigned sequence of points?
5. Can an entire function omit the value $7-2i$ and assume every other value infinitely often?
6. Why are the points 0 and ∞ so often different from all other values?
7. What kind of function has no singularities in the extended plane?
8. How do the singularities of $f(z)$ compare with those of $1/f(z)$? With those of $f(1/z)$? With those of $1/f(1/z)$?
9. Can a pole be a non-isolated singularity?
10. How does Picard's great theorem compare with Picard's theorem stated in Section 8.2?
11. If $f(z)$ has a pole of order k at z_0, what are the most and least number of terms for the principal part of the Laurent expansion?
12. When can one have an accumulation of singularities?
13. If a function has an absolute value greater than 1 near an isolated singularity, what kind of singularity can it be?
14. If $f(z)$ has an isolated singularity at $z=0$, what can you say about f if $\lim_{z\to 0}|z|^{2/3}|f(z)|=2$?

Exercises 9.22.

1. Suppose that $f(z)$ has an isolated singularity at $z=z_0$, and that $\lim_{z\to z_0}(z-z_0)^\alpha f(z)=M\neq 0,\infty$. Prove that α must be an integer.

2. If $f(z)$ is analytic in a deleted neighborhood of the origin and
$$\lim_{z\to 0} |zf(z)| = 0,$$
show that the origin is a removable singularity of $f(z)$.
3. Show that $\tan z$ does not assume the value $\pm i$. Does this contradict Picard's theorem?
4. Find all singularities for the following functions, and describe their nature.

(a) $\tan z$ (b) $\dfrac{1}{e^{1/z}+1}$ (c) $\dfrac{1}{z^2(e^z-1)}$

(d) $\dfrac{1}{\sin z - \cos z}$ (e) $e^{z+1/z}$ (f) $\dfrac{1}{\cos(1/z)}$

(g) $\dfrac{(z-1)^{1/2}}{z+1}$ (h) $\dfrac{\sin^4 z}{z^4} + \cos(3z)$ (i) $\dfrac{z}{e^z-1}$.

5. Discuss the singularities of
$$f(z) = \frac{z^3(z^2-1)(z-2)^2}{\sin^2(\pi z)} e^{1/z^2}.$$
Classify which of these are poles, removable singularities and essential singularity.
6. Describe the singularity at $z=\infty$ for the following functions.

(a) $\dfrac{2z^2+1}{3z^2-10}$ (b) $\dfrac{z^2}{z+1}$ (c) $\dfrac{z^2+10}{e^z}$

(d) $\dfrac{e^z}{z^2+10}$ (e) $\tan z - z$ (f) $\dfrac{1}{z} + \sin z$.

7. Given arbitrary distinct complex numbers z_0, z_1 and z_2, construct a function $f(z)$ having a removable singularity at $z=z_0$, a pole of order k at $z=z_1$, and an essential singularity at $z=z_2$.
8. Show that $f(z)$ has no singularities in the extended plane other than poles if and only if $f(z)$ is a rational function (quotient of two polynomials).
9. If $f(z)$ has poles at a sequence of points $\{z_n\}$, and $z_n \to z_0$, show that $f(z)$ does not have a pole at $z=z_0$. Illustrate this fact by a concrete example.
10. Suppose $f(z)$ has a pole of order m at $z=z_0$, and $P(z)$ is polynomial of degree n. Show that $P(f(z))$ has a pole of order mn at $z=z_0$.
11. Determine the order of the pole at $z=0$ for

(i) $f(z) = \dfrac{z}{\sin z - z + z^3/3!}$ (ii) $f(z) = \dfrac{z}{(\sin z - z + z^3/3!)^2}$.

12. Use "long division" method (or other method) to find the principal part in the Laurent series of $f(z) = 1/(1-\cos z)$ about $z=0$.

13. Let $f(z)$ be analytic in the disk $|z| < R$ $(R > 1)$ except for a simple pole at a point z_0, $|z_0| = 1$. Consider the expansion $f(z) = a_0 + a_1 z + a_2 z^2 + \cdots$, and show that $\lim_{n\to\infty}(a_n/a_{n+1}) = z_0$.
14. Consider, in a neighborhood of the origin, the various determinations of $(1+z)^{1/z}$.
    (a) Show that one of them is analytic in $|z| < 1$, and denote it by $f_0(z)$.
    (b) Determine a_0, a_1 and a_2 in the expansion $f_0(z) = a_0 + a_1 z + a_2 z^2 + \cdots$.
    (c) Let $f(z)$ be a determination of $(1+z)^{1/z}$ other than $f_0(z)$. Find the nature of $g(z) = f(z)/f_0(z)$, and give its Laurent expansion for $|z| > 0$.

9.3 Evaluation of Real Integrals

If $f(z)$ is analytic in a deleted neighborhood of z_0, then by Laurent's theorem we may write

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \quad (0 < |z - z_0| < \delta),$$

where

$$a_n = \frac{1}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^{n+1}}\, dz \quad (n \in \mathbb{Z}).$$

Here C is any simple closed contour enclosing z_0 and contained in the neighborhood. In particular,

$$a_{-1} = \frac{1}{2\pi i} \int_C f(z)\, dz, \quad \text{i.e.,} \quad \int_C f(z)\, dz = 2\pi i a_{-1}. \tag{9.13}$$

Therefore, by hook or crook, we should be able to compute a_{-1}. The coefficient a_{-1} is called the *residue* of $f(z)$ at z_0 and is denoted by

$$\mathrm{Res}\,[f(z); a].$$

Equation (9.13) says that evaluating a certain integral of $f(z)$ around C that encloses no other singularity other than z_0 is akin to determining a certain coefficient in Laurent series.

Examples 9.23. (i) As $e^{1/z} = 1 + 1/z + 1/(2!z^2) + \cdots$ for $|z| > 0$, $\mathrm{Res}\,[e^{1/z}; 0] = 1$ and so, we have

$$\int_{|z|=1} e^{1/z}\, dz = 2\pi i.$$

More generally,

$$\int_C \exp\left(\frac{1}{z^k}\right) dz = \begin{cases} 0 & \text{if } k \neq 1 \\ 2\pi i & \text{if } k = 1 \end{cases}, \quad k \in \mathbb{Z},$$

where C is a simple closed contour enclosing the origin.

(ii) As

$$\sin\left(\frac{1}{z^2}\right) = \frac{1}{z^2} - \frac{1}{3!}\left(\frac{1}{z^2}\right)^3 + \frac{1}{5!}\left(\frac{1}{z^2}\right)^5 - \cdots \quad \text{for } |z| > 0,$$

we have $\mathrm{Res}\,[\sin(1/z^2); 0] = 0$, and so

$$\int_{|z|=1} \sin\frac{1}{z^2}\, dz = 0.$$

More generally,

$$\int_C \sin\left(\frac{1}{z^k}\right) dz = \begin{cases} 0 & \text{if } k \neq 1 \\ 2\pi i & \text{if } k = 1 \end{cases}, \quad k \in \mathbb{Z},$$

where C is a simple closed contour enclosing the origin.

(iii) For $z \neq 0$, we have $\mathrm{Res}\,[z^2 \sin(1/z^2); 0] = -1/6$. Therefore

$$\int_{|z|=1} z^2 \sin\frac{1}{z}\, dz = 2\pi i \left(-\frac{1}{6}\right) = -\frac{\pi i}{3},$$

where C is a simple closed contour enclosing the origin.

(iv) To evaluate $I = \displaystyle\int_{|z|=\pi} \frac{e^z - 1}{1 - \cos z}\, dz$, we consider

$$f(z) = \frac{e^z - 1}{1 - \cos z} = \frac{e^z - 1}{2\sin^2(z/2)}$$

and note that

$$\lim_{z\to 0} z\left(\frac{e^z - 1}{2\sin^2(z/2)}\right) = \lim_{z\to 0} 2\left(\frac{z/2}{\sin(z/2)}\right)^2 \frac{e^z - 1}{z} = 2.$$

Thus, $z = 0$ is a simple pole for $f(z)$. Note also that $f(z)$ has no other singularity inside the circle $|z| = \pi$. Hence, $I = 4\pi i$.

(v) To evaluate $I = \displaystyle\int_{|z|=\pi} \frac{\sin z}{1 - \cos z}\, dz$, we may rewrite the integral as

$$\begin{aligned} I &= \int_{|z|=\pi} \frac{2\sin(z/2)\cos(z/2)}{2\sin^2(z/2)}\, dz \\ &= \int_{|z|=\pi} \frac{\cos(z/2)}{\sin(z/2)}\, dz \\ &= 2\pi i \mathrm{Res}\left[\frac{\cos(z/2)}{\sin(z/2)}; 0\right] \\ &= 2\pi i \lim_{z\to 0} z \frac{\cos(z/2)}{\sin(z/2)} \\ &= 4\pi i. \end{aligned}$$

(vi) To evaluate $I = \int_0^{2\pi} e^{e^{i\theta}-in\theta}\, d\theta$ for $n \in \mathbb{Z}$, we first rewrite it in the form

$$\begin{aligned}
I &= \int_{|z|=1} \frac{e^z}{z^n}\frac{dz}{iz} \quad (z = e^{i\theta}, dz = iz\, d\theta) \\
&= \frac{1}{i}\int_{|z|=1} \frac{e^z}{z^{n+1}}\, dz \\
&= \begin{cases} 0 & \text{if } n = -1, -2, -3, \ldots \text{ (by Cauchy's theorem)} \\ \dfrac{1}{i}\dfrac{2\pi i}{n!} & \text{if } n = 0, 1, 2, \ldots \text{ (by Cauchy's integral formula).} \end{cases}
\end{aligned}$$ ●

Consider now the following generalization of (9.13).

Theorem 9.24. (Residue Theorem) *Suppose $f(z)$ is analytic inside and on a simple closed contour C except for isolated singularities at $z_1, z_2, z_3, \ldots, z_n$ inside C. Then*

$$\int_C f(z)\, dz = 2\pi i \sum_{k=1}^{n} \operatorname{Res}\,[f(z); z_k].$$

Proof. About each singularity z_k construct a circle C_k contained in C and such that $C_j \cap C_k = \emptyset$ when $j \neq k$ (see Figure 9.3). By Cauchy's integral formula for multiply connected domains,

$$\int_C f(z)\, dz = \int_{C_1} f(z)\, dz + \int_{C_2} f(z)\, dz + \cdots + \int_{C_n} f(z)\, dz,$$

where the integration along each interior contour is counterclockwise. Setting $C = C_k$ in (9.13), we see that

$$a_{-1}^{(k)} := \frac{1}{2\pi i}\int_{C_k} f(z)\, dz = \operatorname{Res}\,[f(z); z_k]$$

for each k, and the result follows. ■

As a matter of fact, Cauchy's integral formula is a special case of the residue theorem. To see this, we suppose that $f(z)$ is analytic inside and on a

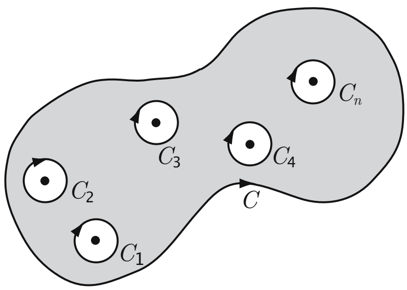


Figure 9.3.

simple closed contour C containing z_0. Then $g(z) = f(z)/(z-z_0)$ has a simple pole at z_0 provided that $f(z_0) \neq 0$. The residue of $g(z)$ at z_0 is given by

$$\mathrm{Res}\,[g(z); z_0] = \lim_{z \to z_0} (z-z_0)g(z) = f(z_0),$$

and so

$$\int_C g(z)\,dz = \int_C \frac{f(z)}{z-z_0}\,dz = 2\pi i f(z_0).$$

Thus, the Cauchy integral formula is a special case of the residue theorem.

Next, suppose $f(z)$ has a pole of order k at $z = z_0$. To find the residue a_{-1} in terms of $f(z)$, by Laurent's series, we write

$$\begin{aligned}(z-z_0)^k f(z) = a_{-k} + a_{-k+1}(z-z_0) + \cdots \\ +a_{-1}(z-z_0)^{k-1} + g(z)(z-z_0)^k,\end{aligned} \tag{9.14}$$

where $g(z)$ is analytic at $z = z_0$. Differentiating (9.14) $k-1$ times and evaluating at $z = z_0$, we get the following result.

Theorem 9.25. (Residue at a pole of order k) *If $f(z)$ has a pole of order k at $z = z_0$, then*

$$\mathrm{Res}\,[f(z); z_0] = \frac{1}{(k-1)!} \lim_{z \to z_0} \frac{d^{k-1}}{dz^{k-1}} (z-z_0)^k f(z). \tag{9.15}$$

In particular, if f has a simple pole at z_0, then

$$\mathrm{Res}\,[f(z); z_0] = \lim_{z \to z_0} (z-z_0) f(z).$$

The following special case is particularly useful.

Theorem 9.26. (Residue at a simple pole) *Let $f(z)$ and $g(z)$ be analytic at z_0. If $g(z)$ has a simple pole at z_0 and $f(z_0) \neq 0$. Then, we have*

$$\mathrm{Res}\left[\frac{f(z)}{g(z)}; z_0\right] = \frac{f(z_0)}{g'(z_0)} \quad \textit{and} \quad \mathrm{Res}\left[\frac{1}{g(z)}; z_0\right] = \frac{1}{g'(z_0)}.$$

Proof. By hypothesis $f(z)/g(z)$ has a simple pole at z_0. Consequently (as $g(z_0) = 0$ and $g'(z_0) \neq 0$), by Theorem 9.25

$$\mathrm{Res}\left[\frac{f(z)}{g(z)}; z_0\right] = \lim_{z \to z_0} \left(\frac{z-z_0}{g(z)-g(z_0)}\right) f(z) = \frac{f(z_0)}{g'(z_0)}. \qquad \blacksquare$$

To illustrate the use of (9.15), we provide a couple of more examples.

Examples 9.27. (i) To evaluate $I = \int_{|z|=1} |z-a|^{-4}\,|dz|$ for $a > 1$, we may first rewrite

$$\begin{aligned}
\frac{|dz|}{|z-a|^4} &= \frac{d\theta}{(z-a)^2(\overline{z}-a)^2} \quad (z=e^{i\theta} \Rightarrow |dz| = |iz\,d\theta| = d\,\theta) \\
&= \frac{1}{(z-a)^2(1/z-a)^2}\,\frac{dz}{iz} \\
&= \frac{z^2}{(z-a)^2(1-az)^2}\,\frac{dz}{iz} \\
&= \frac{z/(z-a)^2}{ia^2(z-1/a)^2}\,dz,
\end{aligned}$$

and therefore, since $a > 1$,

$$\begin{aligned}
I &= \frac{1}{ia^2}\int_{|z|=1} \frac{z/(z-a)^2}{(z-1/a)^2}\,dz \\
&= \frac{1}{ia^2}\left[2\pi i\,\frac{d}{dz}\left(\frac{z}{(z-a)^2}\right)\bigg|_{z=1/a}\right] \\
&= 2\pi\left(\frac{a^2+1}{(a^2-1)^3}\right).
\end{aligned}$$

(ii) If $f(z) = (z^2+a^2)^{-n}$ for some $a > 0$ and $n \in \mathbb{N}$, then the singularities of $f(z)$ are given by

$$z^2 + a^2 = 0, \quad \text{i.e.,} \quad z = \pm ia.$$

Clearly, $z = \pm ia$ are poles of order n for $f(z)$. If $n = 1$, then

$$\text{Res}[f(z); ia] = \frac{1}{2ia}.$$

For $n > 1$, the residue is given by

$$\begin{aligned}
\text{Res}\,[f(z); ia] &= \frac{1}{(n-1)!}\lim_{z\to ia}\frac{d^{n-1}}{dz^{n-1}}((z-ia)^n f(z)) \\
&= \frac{1}{(n-1)!}\lim_{z\to ia}\frac{d^{n-1}}{dz^{n-1}}\left(\frac{1}{(z+ia)^n}\right) \\
&= \frac{1}{(n-1)!}\lim_{z\to ia}\left[\frac{-n(-n-1)\,\cdots\,(-n-(n-2))}{(z+ia)^{n+n-1}}\right] \\
&= \frac{1}{(n-1)!}\left[\frac{(-1)^{n-1}n(n+1)\,\cdots\,(2n-2)}{(2ia)^{n+n-1}}\right] \\
&= \frac{(2n-2)!}{((n-1)!)^2}\,\frac{i^{2n-2-(2n-1)}}{(2a)^{2n-1}} \\
&= -\frac{(2n-2)!}{((n-1)!)^2(2a)^{2n-1}}.
\end{aligned}$$

●

Suppose that $z = \infty$ is an isolated singularity of $f(z)$. Then $f(z)$ is analytic in a deleted neighborhood of $z = \infty$ and so, by Laurent's theorem, we may write

$$f(z) = \sum_{n=-\infty}^{\infty} a_n z^n \quad (\delta < |z| < \infty),$$

for some $\delta > 0$. Choose $R > \delta$ and let γ be the circle of radius R centered at 0, which is traversed in the *clockwise direction*, so that the point at infinity is to the left as in the case of finite isolated singularity. Note that $\int_\gamma z^n\,dz = 0$ for $n \neq -1$ and $\int_\gamma z^{-1}\,dz = -2\pi i$. Thus, because of the uniform convergence on $|z| = R$, we have

$$\frac{1}{2\pi i}\int_\gamma f(z)\,dz = \frac{1}{2\pi i}\sum_{n=-\infty}^{\infty} a_n \int_\gamma z^n\,dz = -a_{-1}.$$

Therefore, we define the residue of $f(z)$ at $z = \infty$ as

$$\text{Res}\,[f(z);\infty] = \frac{1}{2\pi i}\int_\gamma f(z)\,dz = -\frac{1}{2\pi i}\int_{|z|=R} f(z)\,dz = -a_{-1}$$

where $R > \delta$. Also, as $f(z) = \sum_{n=-\infty}^{\infty} a_n z^n$ is analytic for $|z| > R$ iff $f(1/z) = \sum_{n=-\infty}^{\infty} a_n z^{-n}$ is analytic for $0 < |z| < 1/R$, we have

$$\begin{aligned} a_{-1} &= \text{coefficient of } 1/z \text{ in } \frac{1}{z^2} f\left(\frac{1}{z}\right) = \sum_{n=-\infty}^{\infty} \frac{a_n}{z^{n+2}}, \quad 0 < |z| < 1/R, \\ &= \text{Res}\left[\frac{1}{z^2} f\left(\frac{1}{z}\right);0\right] \end{aligned}$$

and hence,

$$\text{Res}\,[f(z);\infty] = -\text{Res}\left[\frac{1}{z^2} f\left(\frac{1}{z}\right);0\right].$$

For instance, if $f(z) = 1 - 1/z$ for $0 < |z| < \infty$, then

$$g(z) = f(1/z) = 1 - z \;\text{ and }\; (1/z^2)f(1/z) = z^{-2} - z^{-1},$$

showing that $g(z)$ has a removable singularity at the origin. In other words, $f(z)$ has a removable singularity at the point at infinity, and $\text{Res}\,[f(z);\infty] = 1$. Note also that $z = 0$ is the only singularity of $f(z)$ in $\mathbb{C}$ and is a simple pole with $\text{Res}\,[f(z);0] = -1$. Thus,

$$\text{Res}\,[f(z);0] + \text{Res}\,[f(z);\infty] = 0$$

which is a demonstration for the following result.

Theorem 9.28. (Residue Theorem for $\mathbb{C}_\infty$) *Suppose $f(z)$ is analytic in $\mathbb{C}_\infty$ except for isolated singularities at $z_1, z_2, z_3, \dots, z_n, \infty$. Then the sum of its residues (including the point at infinity) is zero. That is,*

$$\mathrm{Res}\,[f(z);\infty] + \sum_{k=1}^{n} \mathrm{Res}\,[f(z);z_k] = 0.$$

Proof. Choose R large enough so that all the isolated singularities in $\mathbb{C}$ are in $|z|<R$. Then, by Theorem 9.24,

$$\frac{1}{2\pi i}\int_{|z|=R} f(z)\,dz = \sum_{k=1}^{n} \mathrm{Res}\,[f(z);z_k].$$

But the integral on the left is $-\mathrm{Res}\,[f(z);\infty]$, and the result follows. ■

Example 9.29. Consider the evaluation of the integral

$$I = \frac{1}{2\pi i}\int_{|z|=2} f(z)\,dz, \quad f(z) = \frac{1}{(z-3)(z^n-1)} \ (n\in\mathbb{N}).$$

By Residue theorem, $I = \sum_{k=1}^{n} \mathrm{Res}\,[f(z);z_k]$ where z_k's are nothing but the nth roots of unity. However, by the residue theorem for the extended complex plane, we must have

$$\sum_{k=1}^{n} \mathrm{Res}\,[f(z);z_k] = -\{\mathrm{Res}\,[f(z);3] + \mathrm{Res}\,[f(z);\infty]\}.$$

We note that $\mathrm{Res}\,[f(z);3] = \lim_{z\to 3}(z-3)f(z) = (3^n-1)^{-1}$ and

$$\mathrm{Res}\,[f(z);\infty] = -\mathrm{Res}\left[\frac{1}{z^2}f\left(\frac{1}{z}\right);0\right] = -\mathrm{Res}\left[\frac{z^{n-1}}{(1-3z)(1-z^n)};0\right] = 0.$$

Hence, $I = -1/(3^n-1)$. ●

Example 9.30. We illustrate Theorem 9.28 by finding residues at all singularities of

$$f(z) = \frac{z^n e^{1/z}}{1+z}, \ n\in\mathbb{N}.$$

This function has a simple pole at $z=-1$ and has an essential singularity at $z=0$. Therefore, $\mathrm{Res}\,[f(z);-1] = (-1)^n/e$. If we let $w = z^{-1}$ we obtain

$$f(z) = f(1/w) = \frac{e^w}{w^{n-1}(1+w)}, \quad 0<|w|<1. \tag{9.16}$$

This implies that $z=\infty$ is a pole of order $n-1$ for $f(z)$. Since $z=0$ is an essential singularity of $f(z)$, we must rely on our ability to find the Laurent series expansion of $f(z)$ around zero. Thus we form

$$f(z) = \left(\sum_{k=0}^{\infty}(-1)^k z^{n+k}\right)\left(\sum_{m=0}^{\infty}\frac{z^{-m}}{m!}\right), \quad 0 < |z| < 1.$$

Collecting the terms involving $1/z$ (use Cauchy product of two convergent series), we have

$$\operatorname{Res}[f(z);0] = a_{-1} = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(n+k)!}.$$

Next we determine the residue at $z=\infty$. For this, using (9.16), we write

$$\begin{aligned} F(w) &= \frac{f(1/w)}{w^2} \\ &= \frac{e^w}{w^{n+1}(1+w)} \\ &= w^{-n-1}\left[\sum_{k=0}^{\infty}(-1)^k w^k\right]\left[\sum_{m=0}^{\infty}\frac{w^m}{m!}\right], \quad 0 < |w| < 1. \end{aligned}$$

Again collecting the terms involving $1/w$ (again use Cauchy product of two convergent series), we have

$$\operatorname{Res}[F(w);0] = \frac{(-1)^n}{0!} + \frac{(-1)^{n-1}}{1!} + \cdots + \frac{(-1)^0}{n!} = \sum_{k=0}^{n}\frac{(-1)^{n-k}}{k!}.$$

Therefore, by the definition of the residue at ∞, we find that

$$\operatorname{Res}[f(z);\infty] = -\operatorname{Res}[F(w);0] = -\sum_{k=0}^{n}\frac{(-1)^{n-k}}{k!}.$$

We see that $\operatorname{Res}[f(z);0] + \operatorname{Res}[f(z);-1] + \operatorname{Res}[f(z);\infty] = 0$. •

Armed with several ways to determine residues, we turn now to an important application, that of evaluating a real integral by integrating a complex function along a simple closed contour. The usual method involves a complex function that is real on the real axis. Then a real interval is one of the smooth curves that make up the contour along which we integrate. Recall that the improper integral $\int_a^{\infty} f(x)\,dx$ is defined to be $\lim_{R\to\infty}\int_a^R f(x)\,dx$ if this limit exists.

Example 9.31. We wish to use contour integration to show that

$$\int_{-\infty}^{\infty}\frac{dx}{x^2+1} = \pi.$$

The complex function $f(z) = 1/(z^2+1)$ has singularities at $z = \pm i$. Let C be the contour consisting of the real axis from $-R$ to R $(R > 1)$ followed by the

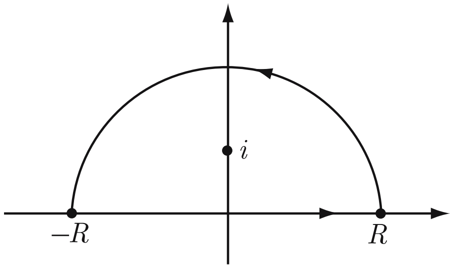


Figure 9.4.

semicircle in the upper half-plane (see Figure 9.4). Then the only singularity of $f(z)$ inside C is at $z = i$, and its residue is

$$\lim_{z \to i} (z - i) f(z) = \frac{1}{2i}.$$

Hence

$$\int_C \frac{dz}{z^2 + 1} = 2\pi i \frac{1}{2i} = \pi \tag{9.17}$$

and the value of this integral is independent of R, $R > 1$. Also,

$$\int_C \frac{dz}{z^2 + 1} = \int_{-R}^{R} \frac{dx}{x^2 + 1} + \int_0^\pi \frac{iRe^{i\theta}}{R^2 e^{2i\theta} + 1}\, d\theta. \tag{9.18}$$

Observe that

$$\left| \int_0^\pi \frac{iRe^{i\theta}}{R^2 e^{2i\theta} + 1}\, d\theta \right| \le \int_0^\pi \frac{R}{R^2 - 1}\, d\theta = \frac{\pi R}{R^2 - 1}. \tag{9.19}$$

In view of (9.19), the second integral on the right of (9.18) approaches 0 as $R \to \infty$. Thus

$$\lim_{R \to \infty} \int_C \frac{dz}{z^2 + 1} = \int_{-\infty}^{\infty} \frac{dx}{x^2 + 1},$$

and the result follows from (9.17). ●

In evaluating a real integral by contour integration, an appropriate complex function and an appropriate contour must be chosen. In Example 9.31, the choice of the complex function was easy. That this is not always the case will be seen shortly. The reader should verify that the desired result in Example 9.31 could also have been obtained using the contour consisting of the real axis from $-R$ to R followed by the semicircle in the *lower* half-plane. The technique of Example 9.31 can be adopted to evaluate integrals of the form

$$I = \int_{-\infty}^{\infty} \frac{p(x)}{q(x)}\, dx,$$

where $p(x)$ and $q(x)$ are polynomials such that

(i) $q(x) \neq 0$ for $x \in \mathbb{R}$
(ii) $p(x)$ and $q(x)$ have real coefficients
(iii) $\deg q(x) \geq \deg p(x) + 2$.

In view of these assumptions, if we proceed exactly as in Example 9.31, it follows that the integral over the semicircular contour

$$\Gamma_R = \{z = Re^{i\theta} : 0 \leq \theta \leq \pi\}$$

in the upper half-plane approaches zero as $R \to \infty$. Consequently, the value of the integral I is $2\pi i$ times the sum of the residues evaluated at those singularities which lie in the upper half-plane.

The same contour may be used to evaluate integrals of the form

$$\int_{-\infty}^{\infty} \frac{p(x)}{q(x)} \cos(ax)\, dx \quad \text{and} \quad \int_{-\infty}^{\infty} \frac{p(x)}{q(x)} \sin(ax)\, dx \quad (a > 0),$$

where p and q are as above. One may weaken the condition (iii) described above by replacing it by

$$\deg q(x) \geq \deg p(x) + 1,$$

with a slightly different argument (as we shall see in some examples below). The integrals of this form are encountered in applications of Fourier analysis, and so they are often referred to as a special case of Fourier integrals.

Clearly, to use the semicircular contour, we cannot start with

$$f(z) = \frac{p(z)}{q(z)} \cos(az) \quad \text{or} \quad \frac{p(z)}{q(z)} \sin(az) \quad (a > 0)$$

because both $\cos z$ and $\sin z$ grow faster than polynomials along the imaginary axis. The trick is to start with

$$f(z) = \frac{p(z)}{q(z)} e^{iaz}$$

and recover the cosine and sine integrals at the end by taking real and imaginary parts, respectively. Note that for $\operatorname{Im} z \geq 0$ and $a > 0$,

$$|e^{iaz}| = |e^{ia(x+iy)}| = e^{-ay} \leq e^0 = 1,$$

so that e^{iaz} is bounded by 1 for all z in the upper half-plane $\{z : \operatorname{Im} z \geq 0\}$. Note that for $a < 0$, e^{iaz} is bounded on the lower half-plane $\{z : \operatorname{Im} z \leq 0\}$ but not on the upper half-plane. In this situation either one has to choose the lower half-plane or start with

$$\frac{p(z)}{q(z)} e^{-iaz}.$$

As another example, we next show that

$$\int_{-\infty}^{\infty} \frac{\cos(ax)}{x^2+m^2}\,dx = \frac{\pi e^{-a}}{m} \quad \text{for } a, m > 0.$$

Consider $f(z) = e^{iaz}/(z^2+m^2)$ and C is same contour as above, namely, C is the boundary of the semi-disk in the upper half-plane bounded by the interval $[-R, R]$ on the real axis and the semicircular contour Γ_R of radius R (large enough to enclose im inside C) in the upper half-plane. Note that $f(z)$ has only one simple pole inside C at $z = im$ with

$$\operatorname{Res}\,[f(z); im] = \lim_{z\to im}(z-im)\frac{e^{iaz}}{z^2+m^2} = \frac{e^{ia(im)}}{2im} = \frac{e^{-a}}{2im}.$$

As usual,

$$\int_C f(z)\,dz = \int_{-R}^{R} f(x)\,dx + \int_{\Gamma_R} f(z)\,dz = 2\pi i\left(\frac{e^{-a}}{2im}\right) = \frac{\pi e^{-a}}{m}. \tag{9.20}$$

As $|e^{iaz}| = e^{-ay} \le e^0 = 1$ for $\operatorname{Im} z = y \ge 0$, the ML-estimate yields

$$\left|\int_{\Gamma_R} f(z)\,dz\right| = \left|\int_{\Gamma_R} \frac{e^{iaz}}{z^2+m^2}\,dz\right| \le \frac{\pi R}{R^2-m^2} \to 0 \ \text{ as } \ R\to\infty.$$

Consequently, passing to the limit $R \to \infty$ in (9.20) shows that

$$\int_{-\infty}^{\infty} \frac{e^{iax}}{x^2+m^2}\,dx = \frac{\pi e^{-a}}{m}.$$

Equating real parts, we have

$$\int_{-\infty}^{\infty} \frac{\cos ax}{x^2+m^2}\,dx = \frac{\pi e^{-a}}{m}, \quad \text{i.e.,} \quad \int_0^{\infty} \frac{\cos ax}{x^2+m^2}\,dx = \frac{\pi e^{-a}}{2m}.$$

Note also that if we take the imaginary part of the integral, we get

$$\int_{-\infty}^{\infty} \frac{\sin ax}{x^2+m^2}\,dx = 0 \quad (a, m > 0).$$

For our next example, we need the following result.

Lemma 9.32. *If* $0 < \theta \le \pi/2$, *then* $\sin\theta \ge (2/\pi)\theta$.

Proof. Geometrically, the result is clear because the graph of $\sin\theta$ lies above the line segment connecting $(0,0)$ and $(\pi/2, 1)$.

Alternatively, if we set $f(\theta) = (\sin\theta)/\theta$ then, as $f(\pi/2) = 2/\pi$, it suffices to show that $f(\theta)$ is a decreasing function in the interval $[0, \pi/2]$, where we define $f(0) = \lim_{\theta\to 0} f(\theta) = 1$. An application of the mean-value theorem yields

$$\begin{aligned} f'(\theta) &= \frac{\theta\cos\theta - \sin\theta}{\theta^2} = \frac{\cos\theta - (\sin\theta)/\theta}{\theta} \\ &= \frac{\cos\theta - \cos\xi}{\theta} \quad (0 < \xi < \theta). \end{aligned}$$

Since the cosine is a decreasing function in the interval $[0, \pi/2]$, $f'(\theta) < 0$ and the result follows. ■

Example 9.33. We wish to show that

$$\int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

Our first inclination is to integrate $(\sin z)/z$ along the same contour as in the previous example. This does not work for two reasons. First, $(\sin z)/z$ has a singularity at $z = 0$ and we can not usually integrate along a path that passes through a singularity point. But the singularity is removable; so this difficulty can be overcome. Second, and more important as was indicated earlier, the integral of $(\sin z)/z$ along the semicircle does not approach a finite limit as the radius tends to infinity, because for $z = iR$ one sees that

$$\lim_{R\to\infty} \frac{\sin(iR)}{iR} = \lim_{R\to\infty} \frac{e^{-R} - e^{R}}{2i^2 R} \to \infty \quad \text{as } R \to \infty.$$

We will consider the function e^{iz}/z, whose imaginary part on the real axis is $(\sin x)/x$. Our contour C will consist of the real axis from ϵ to R, the semicircle in the upper half-plane from R to $-R$, the real axis from $-R$ to $-\epsilon$, and the semicircle in the upper half-plane from $-\epsilon$ to ϵ (see Figure 9.5). The function e^{iz}/z is analytic inside and on C, so that

$$\begin{aligned} 0 &= \int_C \frac{e^{iz}}{z}\, dz \\ &= \int_\epsilon^R \frac{e^{ix}}{x}\, dx + \int_0^\pi \frac{e^{iRe^{i\theta}}}{Re^{i\theta}} iRe^{i\theta}\, d\theta + \int_{-R}^{-\epsilon} \frac{e^{ix}}{x}\, dx + \int_\pi^0 \frac{e^{i\epsilon e^{i\theta}}}{\epsilon e^{i\theta}} i\epsilon e^{i\theta}\, d\theta \\ &= \int_\epsilon^R \frac{e^{ix} - e^{-ix}}{x}\, dx + i\int_0^\pi e^{iRe^{i\theta}}\, d\theta - i\int_0^\pi e^{i\epsilon e^{i\theta}}\, d\theta \end{aligned}$$

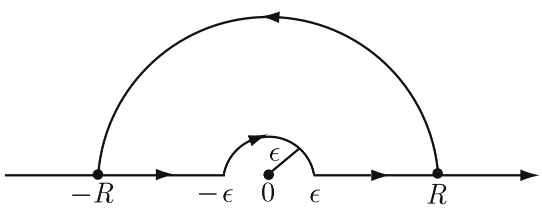


Figure 9.5.

where we have replaced x by $-x$ in the third integral and combined with the first integral. Since $e^{ix} - e^{-ix} = 2i \sin x$, the last equation may be rewritten as

$$0 = 2i \int_{\epsilon}^{R} \frac{\sin x}{x}\, dx + i \int_{0}^{\pi} e^{iRe^{i\theta}}\, d\theta - i \int_{0}^{\pi} e^{i\epsilon e^{i\theta}}\, d\theta. \tag{9.21}$$

We now examine the behavior of the second integral on the left side of (9.21). From the identity $\sin(\pi - \theta) = \sin\theta$ and the lemma, it follows that

$$\begin{aligned} \left| i \int_{0}^{\pi} e^{iRe^{i\theta}}\, d\theta \right| &\leq \int_{0}^{\pi} e^{-R\sin\theta}\, d\theta = 2 \int_{0}^{\pi/2} e^{-R\sin\theta}\, d\theta \\ &\leq 2 \int_{0}^{\pi/2} e^{-(2R/\pi)\theta}\, d\theta \\ &= \frac{\pi}{R}(1 - e^{-R}), \end{aligned}$$

which tends to 0 as R approaches ∞. Hence letting $R \to \infty$ in (9.21) leads to

$$2 \int_{\epsilon}^{\infty} \frac{\sin x}{x}\, dx = \int_{0}^{\pi} e^{i\epsilon e^{i\theta}}\, d\theta. \tag{9.22}$$

For $0 < \epsilon < 1/2$, we expand $e^{i\epsilon e^{i\theta}}$ in a power series to show that

$$|e^{i\epsilon e^{i\theta}} - 1| < 2\epsilon$$

for all θ, $0 < \theta \leq \pi$. We see that

$$\int_{0}^{\pi} e^{i\epsilon e^{i\theta}}\, d\theta = \int_{0}^{\pi} (e^{i\epsilon e^{i\theta}} - 1)\, d\theta + \int_{0}^{\pi} d\theta \to \pi \quad \text{as } \epsilon \to 0.$$

Thus, letting $\epsilon \to 0$ in (9.22), it follows that

$$2 \int_{0}^{\infty} \frac{\sin x}{x}\, dx = \pi$$

and the result follows. The reader should verify that the contour in Figure 9.6 could also have been used to prove the desired result. ●

Let us demonstrate the method by evaluating another integral

$$I = \int_{0}^{\infty} \frac{x \sin(ax)}{x^2 + m^2}\, dx \quad (a, m > 0).$$

Note that the limits of integration in the given integral are not from $-\infty$ to ∞ as required by the method described above. On the other hand, since the integrand is an even function of x,

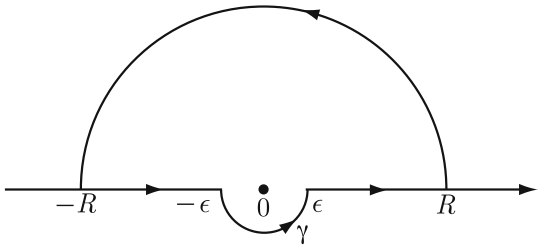


Figure 9.6.

$$I = \frac{1}{2} \int_{-\infty}^{\infty} \frac{x \sin(ax)}{x^2 + m^2}\, dx.$$

Now, we let C be the contour as in Example 9.31 and consider

$$f(z) = \frac{ze^{iaz}}{z^2 + m^2}.$$

Then $f(z)$ has only one simple pole inside C at $z = im$ with

$$\operatorname{Res}\,[f(z); im] = \lim_{z \to im} (z - im) f(z) = \lim_{z \to im} \frac{ze^{iaz}}{z + im} = \frac{e^{-am}}{2}.$$

Thus, for R large enough,

$$\int_C f(z)\, dz = \int_{-R}^{R} f(x)\, dx + \int_{\Gamma_R} f(z)\, dz = 2\pi i \left(\frac{e^{-am}}{2} \right), \tag{9.23}$$

where Γ_R denotes the semicircular contour in the upper half-plane from $-R$ to R. Now, for large R

$$\begin{aligned} \left| \int_{\Gamma_R} f(z)\, dz \right| &= \left| \int_0^{\pi} \frac{Re^{i\theta}}{R^2 e^{2i\theta} + m^2} e^{ia(R\cos\theta + iR\sin\theta)} iRe^{i\theta}\, d\theta \right| \\ &\le \frac{R^2}{R^2 - m^2} \int_0^{\pi} e^{-aR\sin\theta}\, d\theta \\ &\le \frac{R^2}{R^2 - m^2} \left[\frac{\pi}{aR} (1 - e^{-aR}) \right] \end{aligned}$$

which tends to zero as R approaches ∞. Passing to the limit $R \to \infty$ in (9.23) leads to

$$\int_{-\infty}^{\infty} \frac{xe^{iax}}{x^2 + m^2}\, dx = \pi i e^{-am}.$$

Equating the imaginary part gives

$$\int_{-\infty}^{\infty} \frac{x \sin(ax)}{x^2 + m^2}\, dx = \pi e^{-am}$$

and so, $I = \pi e^{-am}/2$.

The next integral is found by methods from calculus. It will be used in conjunction with contour integration to evaluate a different real integral.

Lemma 9.34. $\displaystyle\int_0^\infty e^{-x^2}\,dx = \sqrt{\pi}/2.$

Proof. Set $I = \int_0^R e^{-x^2}\,dx$. Then

$$I^2 = \left(\int_0^R e^{-x^2}\,dx\right)\left(\int_0^R e^{-y^2}\,dy\right) = \int_0^R\int_0^R e^{-(x^2+y^2)}\,dxdy.$$

Here we are integrating along a square S in the first quadrant whose sides have length R. Let C_1 and C_2 be the quarter circles in the first quadrant centered at the origin having radii R and $R\sqrt{2}$, respectively (see Figure 9.7). Evaluating along the circles in polar coordinates, we have

$$\int_0^{\pi/2}\int_0^R e^{-r^2} r\,dr\,d\theta < \int_0^R\int_0^R e^{-(x^2+y^2)}\,dxdy < \int_0^{\pi/2}\int_0^{R\sqrt{2}} e^{-r^2} r\,dr\,d\theta,$$

or

$$\frac{\pi}{4}(1-e^{-R^2}) < \left(\int_0^R e^{-x^2}\,dx\right)^2 < \frac{\pi}{4}(1-e^{-2R^2}).$$

Letting $R\to\infty$, we see that

$$\left(\int_0^\infty e^{-x^2}\,dx\right)^2 = \frac{\pi}{4},$$

and the result follows. ■

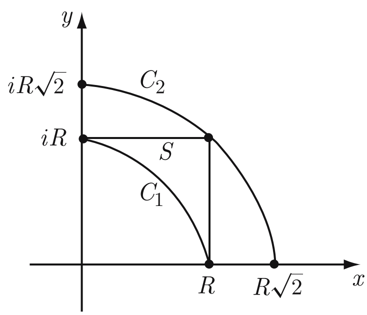


Figure 9.7.

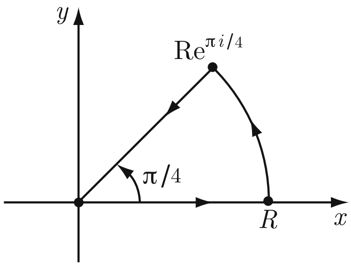


Figure 9.8.

Example 9.35. We wish to show that

$$\int_0^\infty \sin x^2\,dx = \int_0^\infty \cos x^2\,dx = \frac{1}{2}\sqrt{\pi/2}.$$

Note that it is not at all obvious that these integrals even converge.

Let C be the contour consisting of the line segment from 0 to R followed by the arc from R to $Re^{\pi i/4}$ and the line segment from $Re^{\pi i/4}$ to 0 (see Figure 9.8). Since e^{iz^2} is analytic everywhere in $\mathbb{C}$, by Cauchy's theorem, we have

$$\begin{aligned} 0 &= \int_C e^{iz^2}\,dz \\ &= \int_0^R e^{ix^2}\,dx + \int_0^{\pi/4} e^{iR^2e^{2i\theta}} iRe^{i\theta}\,d\theta - \int_0^R e^{i(te^{\pi i/4})^2} e^{\pi i/4}\,dt. \end{aligned}$$

Thus,

$$\int_0^R e^{ix^2}\,dx + \int_0^{\pi/4} e^{iR^2e^{2i\theta}} iRe^{i\theta}\,d\theta = e^{\pi i/4}\int_0^R e^{-t^2}\,dt. \tag{9.24}$$

We now show that the second integral on the left side of (9.24) tends to 0 as R approaches ∞. Note that

$$\begin{aligned} \left|\int_0^{\pi/4} e^{iR^2e^{2i\theta}} iRe^{i\theta}\,d\theta\right| &\le R\int_0^{\pi/4} e^{-R^2\sin 2\theta}\,d\theta \\ &= \frac{R}{2}\int_0^{\pi/2} e^{-R^2\sin\theta}\,d\theta \\ &\le \frac{R}{2}\left[\frac{\pi(1-e^{-R^2})}{2R^2}\right] \\ &\to 0 \quad \text{as } R\to\infty. \end{aligned}$$

In view of this, we let $R\to\infty$ in (9.24) to obtain

$$\int_0^\infty e^{ix^2}\,dx = e^{\pi i/4}\int_0^\infty e^{-t^2}\,dt.$$

Applying Lemma 9.34, we get

$$\int_0^\infty \cos x^2\,dx + i\int_0^\infty \sin x^2\,dx = e^{\pi i/4}\frac{\sqrt{\pi}}{2} = \frac{1}{2}\sqrt{\frac{\pi}{2}} + \frac{i}{2}\sqrt{\frac{\pi}{2}}.$$

The result now follows upon equating real and imaginary parts. •

Our next example involves integrals of trigonometric functions. If z traverses the unit circle $|z| = 1$, then we may parameterize z by

$$z = e^{i\theta} \quad (0 \le \theta \le 2\pi).$$

The identities

$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} = \frac{z - 1/z}{2i}, \quad \text{and} \quad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} = \frac{z + 1/z}{2}$$

enable us to evaluate certain integrals of the form

$$\int_0^{2\pi} g(\sin\theta, \cos\theta)\,d\theta$$

by the residue theorem in the normal way.

To illustrate this, we let a and b real, $|a| > |b|$, and show that

$$I = \int_0^{2\pi} \frac{d\theta}{a + b\cos\theta} = \frac{2\pi}{\sqrt{a^2 - b^2}}. \tag{9.25}$$

The idea is to convert this into a contour integral around the unit circle. First, we observe that there is nothing to prove if $b = 0$. For $b \neq 0$, we may rewrite I as

$$I = \frac{1}{b}\int_0^{2\pi} \frac{d\theta}{a/b + \cos\theta} = \frac{1}{b}\int_0^{2\pi} \frac{d\theta}{\alpha + \cos\theta}$$

where $\alpha = a/b \in \mathbb{R}$ with $|\alpha| > 1$. Thus, it suffices to deal with the case $b = 1$; i.e., to evaluate

$$J = \int_0^{2\pi} \frac{d\theta}{\alpha + \cos\theta}, \quad \text{for } \alpha \text{ real with } |\alpha| > 1.$$

Now, setting $z = e^{i\theta}$, we see that $dz = ie^{i\theta}\,d\theta = iz\,d\theta$. Thus

$$\int_0^{2\pi} \frac{d\theta}{\alpha + \cos\theta} = \int_C \frac{1}{\alpha + (1/2)(z + 1/z)}\frac{dz}{iz} = \frac{2}{i}\int_C f(z)\,dz, \tag{9.26}$$

where C is the unit circle $|z| = 1$ and

$$f(z) = \frac{1}{z^2 + 2\alpha z + 1}.$$

Note that $f(z)$ has simple poles at

$$z_1 = -\alpha + \sqrt{\alpha^2 - 1} \text{ and } z_2 = -\alpha - \sqrt{\alpha^2 - 1}.$$

Note that z_1 lies inside C if $\alpha > 1$ and lies outside C if $\alpha < -1$. A similar observation implies that z_2 lies inside C if $\alpha < -1$ and lies outside if $\alpha > 1$. Now,

$$\text{Res}\,[f(z); z_1] = \lim_{z \to z_1} (z - z_1) f(z) = \lim_{z \to z_1} \frac{z - z_1}{(z - z_1)(z - z_2)} = \frac{1}{z_1 - z_2}$$

and

$$\text{Res}\,[f(z); z_2] = -\frac{1}{z_1 - z_2}.$$

Hence, applying the residue theorem for the last integral in (9.26), we get

$$\int_0^{2\pi} \frac{d\theta}{\alpha + \cos\theta} = \begin{cases} \dfrac{2\pi i}{i\sqrt{\alpha^2 - 1}} & \text{if } \alpha > 1 \\ -\dfrac{2\pi i}{i\sqrt{\alpha^2 - 1}} & \text{if } \alpha < -1 \end{cases} = \begin{cases} \dfrac{2\pi}{\sqrt{\alpha^2 - 1}} & \text{if } \alpha > 1 \\ \dfrac{-2\pi}{\sqrt{\alpha^2 - 1}} & \text{if } \alpha < -1 \end{cases}$$

and hence (9.25) follows. As a consequence, we can easily obtain the following:

(i) $I = \displaystyle\int_0^{\pi} \frac{d\theta}{a^2 + \cos^2\theta} = \int_0^{\pi} \frac{d\theta}{a^2 + \sin^2\theta} = \frac{\pi}{a\sqrt{1 + a^2}}$ for $a > 0$. To do this, we first recall that

$$2\cos^2\theta = 1 + \cos 2\theta \text{ and } 2\sin^2\theta = 1 - \cos 2\theta,$$

and so

$$\int_0^{\pi} \frac{d\theta}{a^2 + \cos^2\theta} = \int_0^{\pi} \frac{2d\theta}{2a^2 + 1 + \cos 2\theta} = \int_0^{2\pi} \frac{d\phi}{\alpha + \cos\phi},$$

where $\alpha = 2a^2 + 1 > 1$. Similarly, we have

$$\int_0^{\pi} \frac{d\theta}{a^2 + \sin^2\theta} = \int_0^{\pi} \frac{2d\theta}{2a^2 + 1 - \cos 2\theta} = -\int_0^{2\pi} \frac{d\phi}{\cos\phi - \alpha}.$$

The desired conclusion follows from (9.25).

(ii) We can also apply (9.25) to show that

$$I = \int_0^{2\pi} \frac{1 + 2\cos\theta}{5 + 4\cos\theta}\, d\theta = 0.$$

To do this, we note that

$$\begin{aligned} I &= \frac{1}{2}\int_0^{2\pi} \frac{5-3+4\cos\theta}{5+4\cos\theta}\,d\theta \\ &= \frac{1}{2}\left[\int_0^{2\pi} d\theta - 3\int_0^{2\pi} \frac{d\theta}{5+4\cos\theta}\right] \\ &= \frac{1}{2}\left[2\pi - 3\left(\frac{2\pi}{\sqrt{25-16}}\right)\right] \\ &= 0. \end{aligned}$$

Let us present another important example of this type. Consider

$$I = \int_0^{2\pi} \frac{\cos n\theta}{\cos\theta + \alpha}\,d\theta, \quad \text{for } \alpha > 1 \text{ and } n \in \mathbb{N}_0.$$

We may rewrite the integral as

$$I = \mathrm{Re}\left[\int_0^{2\pi} \frac{e^{in\theta}}{\cos\theta + \alpha}\,d\theta\right] = \mathrm{Re}\,[J].$$

The substitution $z = e^{i\theta}$ gives

$$J = \int_{|z|=1} \frac{z^n}{\left(\frac{z^2+1}{2z}\right) + \alpha}\,\frac{dz}{iz} = \frac{2}{i}\int_{|z|=1} f(z)\,dz,$$

where

$$f(z) = \frac{z^n}{z^2 + 2\alpha z + 1} = \frac{z^n}{(z-z_1)(z-z_2)}$$

with

$$z_1 = -\alpha + \sqrt{\alpha^2 - 1} \text{ and } z_2 = -\alpha - \sqrt{\alpha^2 - 1}.$$

Since $z_1 z_2 = 1$ and $\alpha > 1$, we have $|z_1| < 1$ and $|z_2| > 1$. It follows that the only singularity of $f(z)$ inside the unit circle $|z| = 1$ is $z = z_1$, which is a simple pole with

$$\mathrm{Res}\,[f(z); z_1] = \lim_{z \to z_1} (z - z_1) f(z) = \lim_{z \to z_1} \frac{z^n}{z - z_2} = \frac{z_1^n}{z_1 - z_2} = \frac{z_1^n}{2\sqrt{\alpha^2 - 1}}.$$

Therefore,

$$J = \frac{2}{i}(2\pi i \mathrm{Res}\,[f(z); z_1]) = 2\pi\left(\frac{(\sqrt{\alpha^2-1} - \alpha)^n}{\sqrt{\alpha^2 - 1}}\right) \quad \text{for } \alpha > 1.$$

Consequently (note that $\mathrm{Im}\,J = 0$), $I = J$.

For instance, for $n = 1$ and $\alpha = 2$, one has

$$\int_0^{2\pi} \frac{\cos\theta}{\cos\theta + 2}\,d\theta = 2\pi\left(1 - \frac{2}{\sqrt{3}}\right).$$

Similarly, for $n = 0$ and $\alpha = 5/4$, one has

$$\int_0^{2\pi} \frac{d\theta}{4\cos\theta + 5} = \frac{1}{4}\int_0^{2\pi} \frac{d\theta}{\cos\theta + 5/4} = \frac{2\pi}{3}.$$

The change of variable $\theta = \pi/2 + \phi$ helps to compute

$$\int_0^{2\pi} \frac{d\theta}{\sin\theta + \alpha} = \int_{\pi/2}^{\pi/2+2\pi} \frac{d\phi}{\cos\phi + \alpha} = \int_0^{2\pi} \frac{d\phi}{\cos\phi + \alpha} \quad \text{for } \alpha > 1.$$

Using the same idea one can compute the integral

$$\int_0^{2\pi} \frac{\cos n\theta}{(\cos\theta + \alpha)^2}\, d\theta \quad \text{for } \alpha > 1.$$

Let us now evaluate

$$I = \int_0^{2\pi} \frac{\cos n\theta}{1 - 2\alpha\cos\theta + \alpha^2}\, d\theta \quad \text{for } -1 < \alpha \neq 0 < 1,\ n \in \mathbb{N}_0.$$

As before

$$I = \mathrm{Re}\left[\int_0^{2\pi} \frac{e^{in\theta}}{1 - 2\alpha\cos\theta + \alpha^2}\, d\theta\right] = \mathrm{Re}\int_0^{2\pi} \frac{e^{in\theta}}{(e^{i\theta} - \alpha)(e^{-i\theta} - \alpha)}\, d\theta = \mathrm{Re}\, J.$$

The usual parameterization $z = e^{i\theta}$ for J gives

$$J = \int_{|z|=1} \frac{z^n}{(z-\alpha)(1/z - \alpha)}\,\frac{dz}{iz} = \frac{1}{i}\int_{|z|=1} f(z)\, dz,$$

where

$$f(z) = \frac{z^n}{(z-\alpha)(1 - z\alpha)}$$

and f has two simple poles at $z = \alpha$ and $z = 1/\alpha$. As $-1 < \alpha \neq 0 < 1$, only $z = \alpha$ lies inside the unit circle with

$$\mathrm{Res}\,[f(z); \alpha] = \lim_{z\to\alpha} \frac{z^n}{1 - \alpha z} = \frac{\alpha^n}{1 - \alpha^2}.$$

Then by the residue theorem

$$J = \frac{1}{i}(2\pi i \mathrm{Res}\,[f(z); \alpha]) = \frac{2\pi\alpha^n}{1 - \alpha^2}.$$

Again note that $\mathrm{Im}\, J = 0$. Consequently,

$$I = \frac{2\pi\alpha^n}{1 - \alpha^2} \quad \text{for } \alpha \in (-1, 1) \setminus \{0\} \text{ and } n \in \mathbb{N}_0.$$

In particular, for $n = 0$, I gives

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{1-\alpha^2}{1-2\alpha\cos\theta+\alpha^2}\,d\theta = 1 \quad \text{for} \quad -1<\alpha<1.$$

Note that the integrand is the Poisson kernel (see Corollary 10.20). Also, we observe that one can also obtain the value of I for $\alpha \in \mathbb{R}$ with $|\alpha|>1$. Finally, we show that

$$I=\int_0^{2\pi}\frac{\sin^2\theta}{a+b\cos\theta}\,d\theta=\frac{2\pi}{b^2}(a-\sqrt{a^2-b^2}) \quad (a,b\in\mathbb{R},\ |a|>|b|>0).$$

To do this, we rewrite

$$I=\frac{1}{b}\int_0^{2\pi}\frac{\sin^2\theta}{\alpha+\cos\theta}\,d\theta=\frac{1}{b}J \quad (|\alpha|=|a/b|>1,\ \alpha\in\mathbb{R})$$

where

$$\begin{aligned} J &= \int_{|z|=1}\left(\frac{z-1/z}{2i}\right)^2\frac{1}{\alpha+(z+1/z)\frac{1}{2}}\,\frac{dz}{iz} \\ &= -\frac{1}{2i}\int_{|z|=1} f(z)\,dz, \quad f(z)=\frac{(z^2-1)^2}{z^2(z^2+2\alpha z+1)}. \end{aligned}$$

Note that f has a double pole at $z=0$ and simple poles at

$$z_1=-\alpha+\sqrt{\alpha^2-1} \text{ and } z_2=-\alpha-\sqrt{\alpha^2-1}=\frac{1}{z_1}.$$

The only singularities which lie inside the unit circle $|z|=1$ are at $z=0$ and at $z=z_1$. Now

$$\begin{aligned} \mathrm{Res}\,[f(z);0] &= \text{ coefficient of } z \text{ in } (z^4-2z^2+1)(z^2+2\alpha z+1)^{-1} \\ &= \text{ coefficient of } z \text{ in } (z^4-2z^2+1)(1-(2\alpha z+z^2)+\cdots) \\ &= -2\alpha \end{aligned}$$

and

$$\begin{aligned} \mathrm{Res}\,[f(z);z_1] &= \lim_{z\to z_1}\frac{(z^2-1)^2}{z^2(z-z_2)} = \frac{(z_1^2-1)^2}{z_1^2(z_1-z_2)} \\ &= (z_1-1/z_1)^2\frac{1}{z_1-z_2} \\ &= \frac{(z_1-z_2)^2}{z_1-z_2} \\ &= z_1-z_2 \\ &= 2\sqrt{\alpha^2-1} \end{aligned}$$

so that $J=-(1/2i)[2\pi i(-2\alpha+2\sqrt{\alpha^2-1})]=2\pi(\alpha-\sqrt{\alpha^2-1})$.

Questions 9.36.

1. If f has a removable singularity at ∞, is $\text{Res}\,[f(z),\infty] = 0$?
2. If $\text{Res}\,[f(z),\infty] = 0$, does f have a removable singularity at $z = \infty$?
3. Since $\int_C \sin(1/z^2)\,dz = 0$ along any simple closed contour containing the origin, why is $\sin(1/z^2)$ not analytic?
4. Is the residue theorem valid for non-isolated singularities?
5. For what kinds of functions will we be able to evaluate complex integrals by means of the residue theorem but not Cauchy's integral formula?
6. In evaluating real integrals by contour integration, what general criteria do we have for choosing the proper complex function and the proper contour?
7. Why does the residue theorem not hold for multiple-valued functions?
8. Why is it easier to evaluate integrals of the form
$$\int_{-\infty}^{\infty} \frac{1}{1+x^{2n}}\,dx$$
than integrals of the form
$$\int_{-\infty}^{\infty} \frac{1}{1+x^{2n+1}}\,dx?$$

Exercises 9.37.

1. Determine the residue at each singularity for the following functions.
(a) $\dfrac{1}{\cos z}$ (b) $\dfrac{z}{(z-1)^2(z-2)}$ (c) $z^n \cos\dfrac{1}{z}$
2. Show that $\displaystyle\int_{|z|=R} |(\sin z)/z|\,|dz| \to \infty$ as $R \to \infty$.
3. Let f be analytic on an open set D, and $f'(a) \neq 0$ for some $a \in D$. Show that
$$\int_C \frac{dz}{f(z)-f(a)} = \frac{2\pi i}{f'(a)}$$
where C is a sufficiently small circle centered at a.
4. Evaluate the following integrals along different simple closed curves not passing through 0 and ± 1.
(i) $\displaystyle\int_C \frac{e^z-1}{z^2(z-1)}\,dz$ (ii) $\displaystyle\int_C \frac{e^z}{z^2(1-z^2)}\,dz.$
5. Evaluate the following integrals.
(a) $\displaystyle\int_{|z|=1/2} \frac{\sin z}{1+z+z^2+\cdots+z^n}\,dz$, where n is the integer nearest your age
(b) $\displaystyle\int_{|z|=5/2} e^{z^2}\pi\cot\pi z\,dz.$

6. For $\phi \in (0, \pi)$ and $n \in \mathbb{N}$, show that

$$\frac{1}{2\pi i} \int_{|z|=2} \frac{z^n}{1 - 2z\cos\phi + z^2}\, dz = \frac{\sin n\phi}{\sin\phi}.$$

7. For each integer n, evaluate

(a) $\displaystyle\int_{|z|=n} \tan \pi z\, dz$ (b) $\displaystyle\int_{|z|=n+1/2} \cot \pi z\, dz.$

8. Evaluate

(i) $\displaystyle\int_{|z|=1} \frac{e^z}{z(2z+1)^2}\, dz$ (ii) $\displaystyle\int_{|z|=2} (2z-1)e^{z/(z-1)}\, dz.$

9. Let $f(z) = z^4 + 6z^2 + 13$. Find the residue of $z^2/f(z)$ at the zeros of $f(z) = 0$ which lie in the upper half-plane $\{w \in \mathbb{C} : \operatorname{Re} w > 0\}$.
10. Using the concept of the residue at the point at infinity, deduce the fundamental theorem of algebra.
11. Use contour integration to evaluate

(a) $\displaystyle\int_0^\infty \frac{dx}{1+x^6}$ (b) $\displaystyle\int_0^\infty \frac{x^2}{1+x^4}\, dx$

(c) $\displaystyle\int_{-\infty}^\infty \frac{dx}{1+x+x^2}$ (d) $\displaystyle\int_{-\infty}^\infty \frac{x}{(x^2+2x+2)^2}\, dx$

(e) $\displaystyle\int_{-\infty}^\infty \frac{\cos ax}{1+x+x^2}\, dx$ (f) $\displaystyle\int_0^\infty \frac{x^2+1}{x^4+1}\, dx$

(g) $\displaystyle\int_0^\infty \frac{x^2}{(x^2+4)^2(x^2+9)}\, dx$ (h) $\displaystyle\int_0^\infty \frac{x^2}{(x^2+1)(x^2+4)}\, dx.$

12. Let C be the rectangle having vertices at 0, R, $R+ic$, ic. By considering the integral $\int_C e^{-z^2}\, dz$, evaluate

$$\int_0^\infty e^{-x^2} \cos(2cx)\, dx.$$

13. (a) By integrating $(e^{iz}-1)/z$ along the contour of Figure 9.4, show that

$$\int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

(b) By integrating $(e^{2iz}-1)/z^2$ along the contour of Figure 9.5, show that

$$\int_0^\infty \frac{\sin^2 x}{x^2}\, dx = \frac{\pi}{2}.$$

(c) By integrating $(1+2iz-e^{iz})/z^2$ along the contour of Figure 9.4, show that

$$\int_0^\infty \frac{\sin^2 x}{x^2}\, dx = \frac{\pi}{2}.$$

14. If $t \neq \pm 1$ is real, show that

$$\int_0^{2\pi} \frac{d\theta}{1 - 2t\cos\theta + t^2} = \frac{2\pi}{|t^2 - 1|}.$$

15. Show that

(a) $\displaystyle\int_0^{2\pi} \frac{\cos\theta}{5 + 4\cos\theta} d\theta = -\frac{\pi}{3}$ (b) $\displaystyle\int_0^{\pi/2} \frac{d\theta}{1 + \sin^2\theta} = \frac{\pi}{2\sqrt{2}}$

(c) $\displaystyle\int_0^{2\pi} \frac{d\theta}{a^2\sin^2\theta + b^2\cos^2\theta} = \frac{2\pi}{ab}$ $(a, b > 0)$.

9.4 Argument Principle

Suppose $f(z)$ is a nonconstant function analytic at z_0 with $f(z_0) = 0$. Then, by Corollary 8.45, there exists a neighborhood of z_0 that contains no other zeros of $f(z)$. Thus we may express $f(z)$ as

$$f(z) = (z - z_0)^k F(z) \quad (k \ \text{ a positive integer}),$$

where $F(z)$ is analytic at z_0 with $F(z_0) \neq 0$. Thus, $F(z) \neq 0$ in the neighborhood of z_0 or on its boundary C. Note that

$$f'(z) = (z - z_0)^{k-1}[kF(z) + (z - z_0)F'(z)],$$

has a zero of order $k - 1$ at z_0 and

$$\frac{f'(z)}{f(z)} = \frac{k}{z - z_0} + \frac{F'(z)}{F(z)} \tag{9.27}$$

so that, at each zero of f of order k, $f'(z)/f(z)$ has a simple pole with residue k. Thus, the residue theorem gives

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\, dz = k,$$

the order of the zero of $f(z)$. The expression $f'(z)/f(z)$ is called the *logarithmic derivative* of $f(z)$ because it is the derivative of $\log f(z)$ at all points where $f(z)$ is analytic and nonzero.

Next suppose that $f(z)$ is analytic inside and on a simple closed contour C with no zeros on C. By Theorem 8.47, $f(z)$ has at most a finite number of zeros inside C. Let the zeros be at $z_1, z_2, \ldots, z_n$ with orders $\alpha_1, \alpha_2, \ldots, \alpha_n$, respectively. Then

$$f(z) = (z - z_1)^{\alpha_1}(z - z_2)^{\alpha_2} \cdots (z - z_n)^{\alpha_n} F(z), \tag{9.28}$$

where $F(z)$ has no zeros inside or on C. Forming the logarithmic derivative in (9.28), we obtain

$$\frac{f'(z)}{f(z)} = \sum_{j=1}^{n} \frac{\alpha_j}{z - z_j} + \frac{F'(z)}{F(z)}. \tag{9.29}$$

An integration of (9.29) leads to

$$\begin{aligned} \frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz &= \frac{1}{2\pi i}\int_C \left(\sum_{j=1}^{n} \frac{\alpha_j}{z-z_j}\right) dz + \frac{1}{2\pi i}\int_C \frac{F'(z)}{F(z)}\,dz \qquad (9.30)\\ &= \sum_{j=1}^{n} \alpha_j \left(\frac{1}{2\pi i}\int_C \frac{dz}{z-z_j}\right). \end{aligned}$$

For each point z_j inside C, construct a circle C_j contained in C having center at z_j and containing no other zero of $f(z)$ (see Figure 9.9). Then for each z_j,

$$\frac{1}{2\pi i}\oint_C \frac{dz}{z-z_j} = \frac{1}{2\pi i}\oint_{C_j} \frac{dz}{z-z_j} = 1. \tag{9.31}$$

An application of (9.31) to (9.30), or a direct application of the residue theorem to (9.29) yields

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz = \sum_{j=1}^{n} \alpha_j. \tag{9.32}$$

Thus, we have

Theorem 9.38. *If f is analytic inside and on a simple closed contour C with no zeros on C, then*

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz = N$$

where N is the number of zeros of $f(z)$ inside C. In determining N, zeros are counted according to their order or multiplicities.

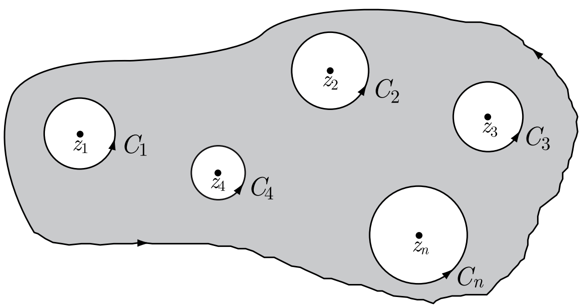


Figure 9.9.

Thus the expression

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz$$

may be viewed as a "counting function" for the zeros of $f(z)$ inside C, where a zero of multiplicity k is counted k times.

In Section 9.2 we have shown that a function $f(z)$ having a pole of order n at z_0 may be expressed as

$$f(z) = \frac{F(z)}{(z-z_0)^n},$$

where $F(z)$ is analytic at $z = z_0$ with $F(z_0) \neq 0$. As can be seen in (9.27), if $f(z)$ has a zero at z_0, then $f'(z)/f(z)$ has a simple pole at z_0.

Equation (9.32) may now be generalized in the following manner:

Theorem 9.39. (Argument Principle) *Let $f(z)$ be analytic inside and on a simple closed contour C except for a finite number of poles inside C, and suppose $f(z) \neq 0$ on C. If N_f and P_f are, respectively, the number of zeros (a zero of order k being counted k times) and poles (again with multiplicity) inside C, then*

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz = N_f - P_f.$$

Proof. Suppose the zeros of $f(z)$ are $z_1, \cdots, z_n$ with multiplicity $\alpha_1, \ldots, \alpha_n$ and the poles of $f(z)$ are $w_1, \ldots, w_m$ with multiplicity $\beta_1, \ldots, \beta_m$. Then $f(z)$ may be written as

$$f(z) = \frac{(z-z_1)^{\alpha_1}\cdots(z-z_n)^{\alpha_n}}{(z-w_1)^{\beta_1}\cdots(z-w_m)^{\beta_m}}F(z),$$

where $F(z)$ is analytic with no zeros or poles inside or on C. Forming the logarithmic derivative, we have

$$\frac{f'(z)}{f(z)} = \sum_{j=1}^{n}\frac{\alpha_j}{z-z_j} - \sum_{j=1}^{m}\frac{\beta_j}{z-w_j} + \frac{F'(z)}{F(z)}. \tag{9.33}$$

Integrating (9.33), we obtain

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz = \sum_{j=1}^{n}\frac{\alpha_j}{2\pi i}\int_C \frac{dz}{z-z_j} - \sum_{j=1}^{m}\frac{\beta_j}{2\pi i}\int_C \frac{dz}{z-w_j}. \tag{9.34}$$

Next enclose each zero and pole of $f(z)$ with disjoint circles containing no other zeros or poles. Then, just as we went from (9.30) to (9.32), so may we go from (9.34) to

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz = \sum_{j=1}^{n}\alpha_j - \sum_{j=1}^{m}\beta_j = N_f - P_f, \tag{9.35}$$

and the proof is complete. ∎

Examples 9.40. Let us now illustrate Theorem 9.39 by simple examples.

(i) Suppose that f is given by

$$f(z) = \frac{(z+1)(z+7)^5(z-i)^2}{(z^2-2z+2)^4(z+i)^8(z-5i)^8}$$

and $C = \{z : |z| = 2\}$. Then, examining the numerator of $f(z)$ shows that inside C, f has a simple zero at $z = -1$, and a zero of order 2 at $z = i$. Therefore, the number N of the zeros of f inside C is

$$N = 1 + 2 = 3.$$

Similarly, as $z^2 - 2z + 2 = (z-1)^2 + 1 = 0$ implies that $z = 1 \pm i$, inside C, f has a pole at $z = 1 + i$ (order 4), $z = 1 - i$ (order 4) and $z = -i$ (order 8). Thus, the number P of the poles of f inside C is

$$P = 4 + 4 + 8 = 16.$$

According to Theorem 9.39,

$$\int_{|z|=2} \frac{f'(z)}{f(z)}\, dz = 2\pi i(N - P) = -26\pi i.$$

(ii) Let us use the Argument principle to evaluate

$$I = \int_C \frac{z+i}{z^2+2iz-4}\, dz, \quad C = \{z : |z+1+i| = 2\}.$$

Note that this integral may be evaluated either by the Cauchy integral formula or the residue theorem. We rewrite I as

$$I = \frac{1}{2}\int_C \frac{f'(z)}{f(z)}\, dz, \quad f(z) = (z+i)^2 - 3.$$

Note that the zeros of f are given by $z = \pm\sqrt{3} - i$. We observe that $-i + \sqrt{3}$ lies outside C, while $z = -\sqrt{3} - i$ lies inside C. Consequently, by Theorem 9.39, we have

$$I = \frac{1}{2}(2\pi i) = \pi i.$$

(iii) We easily see that

$$\int_{|z|=2} \frac{dz}{3z+4} = \frac{1}{3}\int_{|z|=2} \frac{3}{3z+4}\, dz = \frac{1}{3}(2\pi i) \quad (\text{with } f(z) = 3z+4).$$

(iv) To evaluate $I = \int_{|z|=2} \frac{z+2}{z(z+1)}\,dz$, we may rewrite it as

$$I = \int_{|z|=2} \frac{\frac{1}{z^2}+\frac{2}{z^3}}{\frac{1}{z}+\frac{1}{z^2}}\,dz = -\int_{|z|=2} \frac{f'(z)}{f(z)}\,dz,$$

where

$$f(z) = \frac{1}{z} + \frac{1}{z^2} = \frac{z+1}{z^2}.$$

Then, by the Argument principle,

$$I = -2\pi i[N_f - P_f] = -2\pi i[1-2] = 2\pi i.$$

This can be checked by using partial fractions and Cauchy's integral formula:

$$\begin{aligned} I = \int_{|z|=2} \frac{z+1+1}{z(z+1)}\,dz &= \int_{|z|=2} \frac{dz}{z} + \int_{|z|=2} \left(\frac{1}{z} - \frac{1}{z+1}\right) dz \\ &= 2\pi i + 2\pi i - 2\pi i = 2\pi i. \end{aligned}$$ ●

It seems strange indeed that (9.35) is always an integer regardless of the function $f(z)$ or the closed contour C. This phenomenon is based on properties of the logarithm. Suppose that $f(z)$ is analytic and nonzero for all z on a simple closed contour C. Set

$$\log f(z) = \ln |f(z)| + i \arg f(z),$$

where a fixed branch for the logarithm is chosen. Then

$$\int_C \frac{f'(z)}{f(z)}\,dz = \int_C d(\log f(z)) = \ln |f(z)| \Big|_C + i \arg f(z) \Big|_C . \tag{9.36}$$

Since the initial and terminal points of the closed contour C must coincide, $\ln |f(z)| \;|_C = 0$. Hence (9.36) simplifies to

$$\int_C \frac{f'(z)}{f(z)}\,dz = i \arg f(z) \Big|_C . \tag{9.37}$$

Thus the value of the integral depends only on the net change in the argument of $f(z)$ as z traverses the contour C.

Now the image of the simple closed contour C under $f(z)$ is a closed contour C', which need not be simple. We illustrate two cases:

Case 1: If $f(z)$ has no zero inside C, then C' does not surround the origin. Therefore, $\arg f(z)$ returns to its original value as $f(z)$ traverses the contour C' (see Figure 9.10). Let $z_0 \in C$ be mapped to $w_0 \in C'$. As z_0 traverses the contour C once in the positive direction, w_0 traverses C' an integer number of times in the positive or negative direction. However, the number

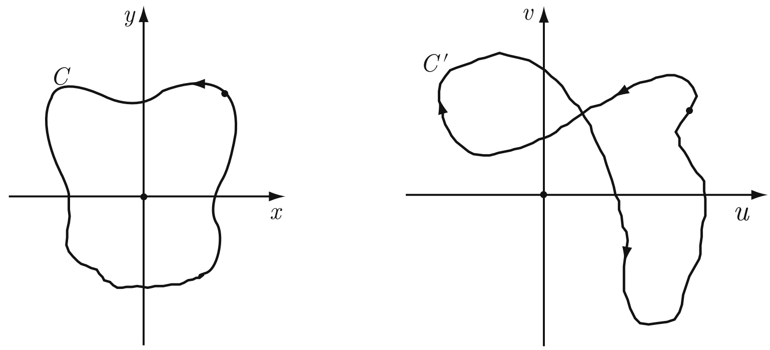


Figure 9.10.

$$\arg f(z)\Big|_{z=z_0} = \arg f(z_0) = \arg w_0$$

does not change as z_0 travels once or several times along C. For example, in Figure 9.10, $\arg w_0$ increases to $\arg A$ (up to the point A) then decreases and when w_0 returns to its initial position, $\arg w_0$ returns to its initial value. This means that the net change in $\arg f(z)$, as z traverses the contour C, is zero. That is,

$$\int_{\Gamma=f(C)} \frac{dw}{w} = \int_C \frac{df(z)}{f(z)}\, dz = \int_C \frac{f'(z)}{f(z)}\, dz = i \arg f(z)\Big|_C = 0.$$

Case 2: If $f(z)$ does have zeros inside C, then C' must wrap around the origin (Why?). Each time that C' winds around the origin (in the positive sense), the argument of $f(z)$ is increased by 2π. In Figure 9.11, we show a simple closed contour C being mapped by $f(z) = z^2$ onto a closed contour C' that twice winds around the origin. When z returns to its initial point on C, $\arg f(z)$ has increased by 4π along C'.

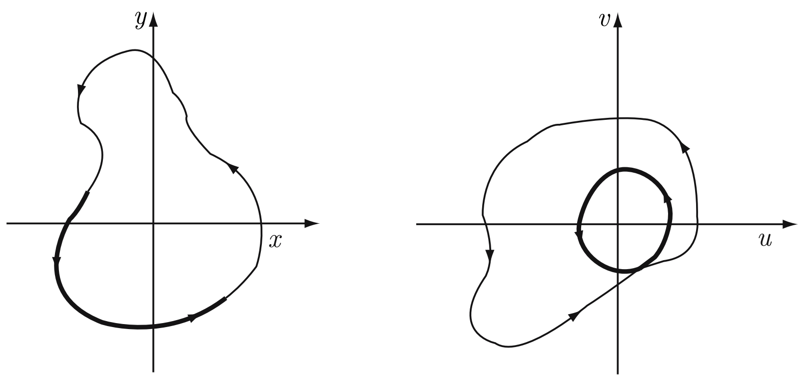


Figure 9.11.

Note that we are not concerned with whether or not the origin is inside the simple closed contour C. The origin only plays a critical role relative to the image curve C'. The net change in $\arg f(z)$ as z traverses a contour C is called the *variation* of $\arg f(z)$ along C, and is denoted by $\Delta_C \arg f(z)$. This notation allows us to write (9.37) as

$$\int_C \frac{f'(z)}{f(z)}\,dz = i\Delta_C \arg f(z).$$

Hence the conclusion of Theorem 9.39 can be expressed as

$$\frac{1}{2\pi}\Delta_C \arg f(z) = N_f - P_f. \tag{9.38}$$

Geometrically, the identity in (9.38) is known as the *Argument Principle.*

Remark 9.41. For each zero of $f(z)$ inside C, the curve C' winds once around the origin in the positive sense, whereas for each pole of $f(z)$ inside C, the curve C' winds once around the origin in the negative sense. To prove this, we need a careful definition of *winding number* (see Ahlfors [A] and Ponnusamy [P1]). ●

The following two lemmas are consequences of the argument principle.

Lemma 9.42. *Suppose $f(z)$ and $g(z)$ are analytic inside and on a simple closed contour C with $f(z)$ and $g(z)$ having no zeros on C. Then*

$$\Delta_C \arg f(z)g(z) = \Delta_C \arg f(z) + \Delta_C \arg g(z).$$

Proof. Let $f(z)$ and $g(z)$ have N_1 and N_2 zeros respectively inside C. Then, by the argument principle,

$$\frac{1}{2\pi}\Delta_C \arg f(z) = N_1 \quad \text{and} \quad \frac{1}{2\pi}\Delta_C \arg f(z) = N_2.$$

But $f(z)g(z)$ has $N_1 + N_2$ zeros inside C. Hence

$$\frac{1}{2\pi}\Delta_C \arg f(z)g(z) = N_1 + N_2,$$

and the proof is complete. ■

Lemma 9.43. *Suppose $h(z)$ is analytic on a simple closed contour C with $|h(z)| < 1$ for all z on C. Then $\Delta_C \arg(1 + h(z)) = 0$.*

Proof. The simple closed contour C is mapped by $w = F(z) = 1+h(z)$ onto a closed contour C' contained in the disk $|w-1| < 1$ (see Figure 9.12). Since this disk is in the right half-plane, we may cut the plane along the negative real axis to obtain a branch (the principle branch) for $\arg F(z)$ as $F(z)$ traverses C'. Thus

$$\frac{1}{2\pi}\Delta_C \arg F(z) = \frac{1}{2\pi}\Delta_C \arg(1 + h(z)) = 0.$$ ■

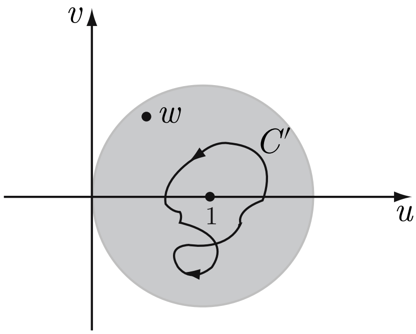


Figure 9.12.

Remark 9.44. The function $F(z) = 1 + h(z)$ need not be analytic for z *inside* C in order for a branch of $\log F(z)$ to exist for z *on* C. In view of the argument principle, we have merely shown that $F(z)$ has the same number of zeros and poles inside C. For example, $h(z) = 1/(2z)$ satisfies the inequality $|h(z)| < 1$ on the unit circle. The function

$$F(z) = 1 + h(z) = \frac{1+2z}{2z}$$

has one simple zero (at $z = -1/2$) and simple pole (at the origin). ●

Theorem 9.45. (Rouché's Theorem) *Suppose $f(z)$ and $g(z)$ are analytic inside and on a simple closed contour C, with $|g(z)| < |f(z)|$ on C. Then $f(z) + g(z)$ has the same number of zeros as $f(z)$ inside C.*

Proof. By hypothesis, $|g(z)| < |f(z)|$ on C, which implies that on C

$$|f(z)| > 0 \quad \text{and} \quad |f(z) + g(z)| \geq |f(z)| - |g(z)| > 0.$$

Thus, f and $f + g$ are analytic inside and on C with f and $f + g$ having no zeros on C. Since f and g are analytic, $P_f = 0 = P_{f+g}$. If we write

$$f + g = f(1 + g/f) := f\phi,$$

then

$$(f+g)'(z) = f'(z)\left(1 + \frac{g(z)}{f(z)}\right) + f(z)\left(1 + \frac{g(z)}{f(z)}\right)'$$

so that

$$\frac{(f+g)'(z)}{(f+g)(z)} = \frac{f'(z)}{f(z)} + \frac{\left(1 + \frac{g(z)}{f(z)}\right)'}{1 + \frac{g(z)}{f(z)}} = \frac{f'(z)}{f(z)} + \frac{\phi'(z)}{\phi(z)}.$$

Let N_f and N_{f+g} denote the number of zeros of f and $f + g$ respectively on the domain which is bounded by C. By the Argument Principle,

$$N_{f+g} - N_f = \frac{1}{2\pi i}\int_C \left(\frac{(f+g)'(z)}{(f+g)(z)} - \frac{f'(z)}{f(z)}\right) dz = \frac{1}{2\pi i}\int_C \frac{\phi'(z)}{\phi(z)}\, dz.$$

Since $|g(z)/f(z)| < 1$ on C, $\phi(z)$ lies in the disc $|w-1| < 1$ for $z \in C$. Taking a branch of the logarithm on the simply connected domain $|w-1| < 1$ (which does contain the origin), we have

$$\frac{\phi'(z)}{\phi(z)} = \frac{d}{dz}\log(\phi(z)).$$

Thus, the integral on the right is zero. That is, $N_{f+g} = N_f$.

Alternatively, by Lemma 9.42,

$$\Delta_C \arg(f(z)+g(z)) = \Delta_C \arg f(z) + \Delta_C \arg\left(1 + \frac{g(z)}{f(z)}\right). \tag{9.39}$$

Since $|g(z)/f(z)| < 1$ on C, Lemma 9.43 may be applied to obtain

$$\Delta_C \arg\left(1 + \frac{g(z)}{f(z)}\right) = 0.$$

Hence (9.39) reduces to

$$\frac{1}{2\pi}\Delta_C \arg(f(z)+g(z)) = \frac{1}{2\pi}\Delta_C \arg f(z). \tag{9.40}$$

Then, by (9.40) and the argument principle, the theorem follows. ■

The proof of Rouché's theorem was "geometric" in character. We now give an "analytic" proof.

Alternate proof of Rouché's Theorem. Let $\{\phi_t(z)\}$ be a family of functions defined by

$$\phi_t(z) = f(z) + tg(z) \quad (0 \le t \le 1).$$

Then, for each t, ϕ_t is analytic inside and on C having no poles inside or on C. Also, since $|f(z)| < |g(z)|$ on C, we have

$$|\phi_t(z)| = |f(z) + tg(z)| \ge |f(z)| - |g(z)| > 0 \qquad \text{for } z \in C$$

and so ϕ_t does not have a zero on C. Since $|f(z)| - |g(z)|$ is continuous on the compact set C, it attains a minimum, say m. Thus for all z on C,

$$|\phi_t(z)| \ge m > 0 \quad (0 \le t \le 1). \tag{9.41}$$

Define

$$h(t) = \frac{1}{2\pi i}\int_C \frac{\phi_t'(z)}{\phi_t(z)}\, dz.$$

Observe that $h(t)$ denotes the number of zeros of $\phi_t(z)$ inside C. We want to show that $h(0) = h(1)$. Now, given any points t_1 and t_2 in $[0,1]$, we have

$$h(t_2) - h(t_1) = \frac{1}{2\pi i} \int_C \left(\frac{f' + t_2 g'}{f + t_2 g} - \frac{f' + t_1 g'}{f' + t_1 g} \right) dz \tag{9.42}$$
$$= \frac{1}{2\pi i} \int_C \frac{(t_2 - t_1)(g'f - f'g)}{(f + t_2 g)(f + t_1 g)}\, dz.$$

Since $f(z)$, $g(z)$, $f'(z)$, and $g'(z)$ are analytic on C, they are bounded, and we may assume that

$$|f(z)|,\ |g(z)|,\ |f'(z)|,\ |g'(z)| \leq M. \tag{9.43}$$

Denote the length of C by L. Then, from (9.41), (9.42), and (9.43), we get

$$|h(t_2) - h(t_1)| \leq \frac{1}{2\pi} \frac{|t_2 - t_1| 2M^2}{m^2} L = K|t_2 - t_1|,$$

where K is constant independent of t_1 and t_2, so that $h(t)$ is a continuous function on $[0, 1]$. Since $h(t)$ is integer-valued, it follows (by the intermediate value theorem) that $h(t)$ is constant on $[0, 1]$. In particular, $h(0) = h(1)$ where $h(0)$ and $h(1)$ are, respectively, the number of zeros of $\phi_0(z) = f(z)$ and the number of zeros of $\phi_1(z) = f(z) + g(z)$ inside C. The proof is complete.

Remark 9.46. In Rouché's theorem, the condition $|g(z)| < |f(z)|$ on C cannot be relaxed to $|g(z)| \leq |f(z)|$. This is seen by setting

$$g(z) = -f(z).$$

Then $f(z) + g(z) \equiv 0$ inside C regardless of the number of zeros of $f(z)$. ●

Corollary 9.47. *Let g be analytic for $|z| \leq 1$ and $|g(z)| < 1$ for $|z| = 1$. Then g has a unique fixed point in $|z| < 1$.*

Proof. For $|z| = 1$, $|g(z)| < |-z| = 1$. By Rouche's theorem, we have $g(z) - z$ has exactly one zero in $|z| < 1$ and the conclusion follows. ■

We ask what happens if we replace $|g(z)| < 1$ with $|g(z)| \leq 1$ in Corollary 9.47? If

$$g(z) = \frac{z - a}{1 - \overline{a}z} \quad (0 < |a| < 1),$$

then g is analytic for $|z| \leq 1$ and $|g(z)| = 1$ for $|z| = 1$. Moreover,

$$g(|z| < 1) \subset \{w : |w| < 1\}$$

and we easily see that

$$g(z) = z \Longleftrightarrow z^2 = a/\overline{a}$$

which shows that, for each a with $0 < |a| < 1$, $g(z) = z$ has no solution in $|z| < 1$.

Next we ask: what happens if we simply assume that f is analytic for $|z| < 1$ and $|f(z)| < 1$ for $|z| < 1$? Of course, $f(z) = z$ shows that every point of $|z| < 1$ is a fixed point. Suppose that $f(z) \not\equiv z$. Can f have more than one fixed point in $|z| < 1$?

To answer this, we suppose that f has two fixed points, say a and b, in $|z| < 1$. Consider

$$\phi(z) = \frac{a - z}{1 - \bar{a}z}.$$

Then, $\phi(a) = 0$ and ϕ is its own inverse. Set $\phi(b) = \alpha$. Then $\alpha \neq 0$, because ϕ is one-to-one. Define $g = \phi \circ f \circ \phi^{-1}$ and see that

$$\begin{aligned} g(0) &= \phi \circ f \circ \phi^{-1}(0) = \phi(f(a)) = \phi(a) = 0 \\ g(\alpha) &= \phi \circ f \circ \phi^{-1}(\alpha) = \phi(f(b)) = \phi(b) = \alpha. \end{aligned}$$

By Schwarz's lemma, $|g(z)| \leq |z|$. But, because equality is attained at an interior point α, we have $g(z) = e^{i\eta}z$ for some real constant η. The condition $g(\alpha) = \alpha$ shows that g must be the identity function. Thus, $f(z) = z$ which is a contradiction.

As a first application of Rouché's theorem, we prove

Theorem 9.48. (Hurwitz's Theorem) *Let $\{f_n(z)\}$ be a sequence of functions analytic inside and on the simple closed contour C, and suppose $\{f_n(z)\}$ converges uniformly to $f(z)$ inside and on C. If $f(z)$ has no zeros on C, then the number of zeros of $f(z)$ inside C is equal to the number of zeros of $f_n(z)$ inside C for sufficiently large n.*

Proof. First note that according to Theorem 8.16, the limit function $f(z)$ is analytic inside and on C. Let $m > 0$ denote the minimum of $|f(z)|$ on C. By the uniform convergence of $\{f_n(z)\}$ on C, we have for $n > N(m)$ that

$$|f_n(z) - f(z)| < m \leq |f(z)|$$

on C. Hence by Rouché's theorem, the number of zeros of $f(z)$ inside C equals the number of zeros of

$$f(z) + (f_n(z) - f(z)) = f_n(z) \quad (n > N). \qquad \blacksquare$$

Rouché's theorem furnishes us with yet another proof for the fundamental theorem of algebra.

Theorem 9.49. (Fundamental Theorem of Algebra) *If*

$$P_n(z) = a_0 + a_1 z + \cdots + a_{n-1}z^{n-1} + a_n z^n \quad (a_n \neq 0)$$

is a polynomial of degree n, then it has n zeros in $\mathbb{C}$.

Proof. Note that for $z \neq 0$,

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

Then for $|z| = r > 1$,

$$\begin{aligned}\left|\frac{P_n(z)}{a_n z^n} - 1\right| &\leq \left[\frac{|a_{n-1}|}{r} + \cdots + \frac{|a_0|}{r^n}\right]\frac{1}{|a_n|} \\ &\leq \frac{|a_0| + |a_1| + \cdots + |a_{n-1}|}{|a_n| r} \\ &< 1 \quad \text{if} \quad r > \max\{(|a_0| + \cdots + |a_{n-1}|)/|a_n|, 1\}.\end{aligned}$$

That is,

$$|P_n(z) - a_n z^n| < |a_n z^n| \quad \text{for} \quad |z| = r.$$

Since $a_n z^n$ has n zeros (all at the origin) inside $|z| = r$, so does the function

$$(P_n(z) - a_n z^n) + a_n z^n = P_n(z). \qquad \blacksquare$$

Remark 9.50. This proof of the fundamental theorem is more satisfying than the previous proofs. First, we get directly that the polynomial has n roots (as opposed to "at least one root"). Second, and more important, we get a bound on the modulus of the roots in terms of the coefficients. We know that all the roots lie in the disk

$$|z| \leq \frac{|a_0| + |a_1| + \cdots + |a_{n-1}|}{|a_n|} \quad (|z| > 1).$$

Rouché's theorem will frequently be an aid in approximating the location of zeros for an analytic function. ●

Example 9.51. Let us use Rouche's theorem to determine the number of zeros of the polynomial $p(z) = z^{10} - 6z^9 - 3z + 1$ inside the unit circle $|z| = 1$. To do this, we set

$$p(z) = f(z) + g(z),$$

where $f(z) = -6z^9 + 1$ and $g(z) = z^{10} - 3z$. Then for $|z| = 1$,

$$|f(z)| = |-6z^9 + 1| \geq |6z^9| - 1 = 6 - 1 = 5$$

and

$$|g(z)| = |z^{10} - 3z| \leq |z|^{10} + 3|z| = 4 < 5 \leq |f(z)|.$$

By Rouche's theorem, $f(z)$ and $p(z)$ have the same number of zeros inside $|z| = 1$. But $f(z)$ has nine zeros inside the unit circle $|z| = 1$. Therefore, $p(z)$ has nine zeros in $|z| < 1$. ●

Example 9.52. Let us show that all five roots of the polynomial

$$P(z) = z^5 + 6z^3 + 2z + 10$$

lie in the annulus $1 < |z| < 3$.

To see this, we let $f(z) = z^5 + 6z^3 + 2z$ and $g(z) = 10$. On $|z| = 1$,

$$|f(z)| \leq |z^5| + 6|z^3| + 2|z| = 9 < |g(z)| = 10.$$

Hence $P(z) = f(z) + g(z)$ has the same number of zeros in $|z| < 1$ as does $g(z)$, namely none. Observe that

$$|P(z)| \geq 10 - |z^5 + 6z^3 + 2z| \geq 10 - |z|^5 - 6|z|^3 - 2|z| = 1 \quad \text{on} \quad |z| = 1$$

and so $P(z)$ can have no zeros on the unit circle $|z| = 1$.

Next let $f(z) = z^5$ and $g(z) = 6z^3 + 2z + 10$. On $|z| = 3$,

$$|g(z)| \leq 6(3^3) + 2(3) + 10 < |f(z)| = 3^5.$$

Thus all the zeros of $P(z)$ must lie in $|z| < 3$, that is, in the annulus $1 < |z| < 3$.

By setting $f(z) = 6z^3$ and $g(z) = z^5 + 2z + 10$, we can further show that three of the roots of $P(z)$ lie in the annulus $1 < |z| < 2$ and, consequently, that the other two lie in the annulus $2 < |z| < 3$. ●

Example 9.53. We easily show that all the roots of

$$z^5 - 3z^2 - 1 = 0$$

lie inside the circle $|z| = 2^{2/3}$ and that two of its roots lie inside the circle $|z| = 3/4$. To do this we first set

$$f(z) = z^5 \quad \text{and} \quad g(z) = -3z^2 - 1.$$

Then, on $|z| = 2^{2/3}$, $|f(z)| = |z|^5 = 2^{10/3}$ and

$$|g(z)| \leq 3|z|^2 + 1 = 3(2^{4/3}) + 1 < 2^{10/3} = |f(z)|,$$

showing that f and $f + g$ have the same number of zeros inside the circle $|z| = 2^{2/3}$. But f has five zeros inside $|z| = 2^{2/3}$. Thus $f + g$ and hence all the roots of given equation lie in $|z| < 2^{2/3}$.

Also, on $|z| = 3/4$, we have $|f(z)| = |z|^5 = (3/4)^5$ and

$$|g(z)| \geq 3|z|^2 - 1 = 3(3/4)^2 - 1 = 11/16 > (3/4)^5 = |f(z)|.$$

It follows that g and $f + g$ have the same number of zeros in $|z| < 3/4$. But g has two zeros at $z = \pm i/\sqrt{3}$ which lies inside the circle $|z| = 3/4$. Hence the given equation has two roots inside $|z| = 3/4$. Consequently, the given equation has three zeros in $3/4 \leq |z| < 2^{2/3}$. ●

Example 9.54. Consider $f(z) = z^2 + 7z + 12 - c$. Then for $|z| \leq 1$,

$$|z^2 + 7z + 12| = |(z + 3)(z + 4)| \geq 2(3) = 6.$$

Therefore, if $|c| < 6$, then $f(z) \neq 0$ in the unit disc $|z < 1$. ●

As a final application of Rouché's theorem, we prove

Theorem 9.55. (Open Mapping Theorem) *A nonconstant analytic function maps open sets onto open sets.*

Proof. Suppose $f(z)$ is analytic at $z = z_0$. We must show that the image of every sufficiently small neighborhood of z_0 in the z plane contains a neighborhood of $w_0 = f(z_0)$ in the w plane. Choose $\delta > 0$ such that the function $f(z) - w_0$ is analytic in the disk $|z - z_0| \leq \delta$ and contains no zeros on the circle $|z - z_0| = \delta$. This is possible in view of Corollary 8.48. Let m be the minimum value of $|f(z) - w_0|$ on the circle $|z - z_0| = \delta$. We will show that the image of the disk $|z - z_0| < \delta$ under $f(z)$ contains the disk $|w - w_0| < m$ (see Figure 9.13).

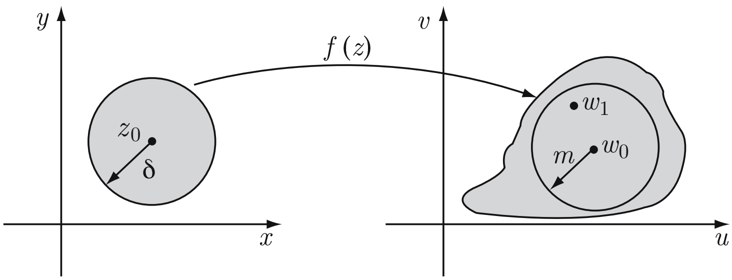


Figure 9.13.

Choose w_1 in the disk $|w - w_0| < m$. Then on the circle $|z - z_0| = \delta$ we have

$$|w_0 - w_1| < m \leq |f(z) - w_0|.$$

Hence by Rouché's theorem,

$$(f(z) - w_0) + (w_0 - w_1) = f(z) - w_1$$

has the same number of zeros in $|z - z_0| < \delta$ as does $f(z) - w_0$. Since $f(z) - w_0$ has at least one zero (at z_0), the function $f(z) - w_1$ has at least one zero. That is, $f(z) = w_1$ at least once. Since w_1 is arbitrary, the image of the disk $|z - z_0| < \delta$ must contain all points in the disk $|w - w_0| < m$. ∎

Corollary 9.56. *A nonconstant analytic function maps a domain onto a domain.*

Proof. Recall that a domain is an open connected set. In view of the theorem, we need only show that an analytic function maps connected sets onto connected sets. But this follows from Exercise 2.46(13) since an analytic function is continuous. ∎

The open mapping theorem provides a short proof of the maximum modulus theorem.

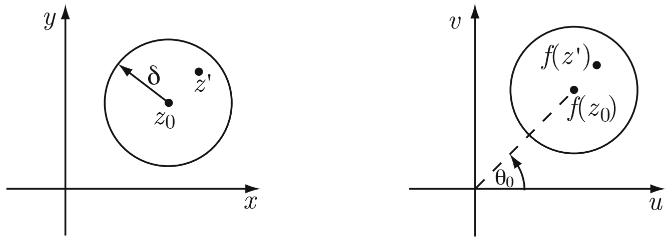


Figure 9.14.

Suppose $f(z)$ is analytic in a domain D and that z_0 is a point in D. If $f(z)$ is not constant, then the image of some disk $|z - z_0| < \delta$ contains a disk $|w - f(z_0)| < m$ in the w-plane. If $f(z_0) = Re^{i\theta_0}$, then to each ϵ, that $f(z') = (R + \epsilon)e^{i\theta_0}$ (see Figure 9.14). Thus,

$$|f(z')| = R + \epsilon > |f(z_0)| = R,$$

so that z_0 is not a maximum for $|f(z)|$.

We end the section with the following corollary which has been proved earlier by a different method (see Theorem 5.37).

Corollary 9.57. *If f is analytic in a domain D and if any one of* $\mathrm{Re}\, f$, $\mathrm{Im}\, f$, $|f|$, *or* $\mathrm{Arg}\, f$ *is constant, then f is also constant.*

Proof. By hypothesis, $f(D)$ would be a subset of either the real axis, or imaginary axis, or a circle or a line with constant argument, respectively. Note that none of them forms an open set. The conclusion follows from the open mapping theorem. ∎

Questions 9.58.

1. Can $f(z)$ be analytic in a deleted neighborhood of z_0 even when the limit $\lim_{z \to z_0} (z - z_0)^n f(z)$ does not exist for any integer n?
2. What is the significance of the constant $2\pi i$?
3. How do the properties of $\Delta_C \arg f(z)$ and $\log f(z)$ compare?
4. Can Rouché's theorem be used to locate the quadrants of zeros for an analytic function?
5. Can Rouché's theorem be extended to the case when there are poles inside the contour?
6. Does a nonconstant continuous function map open sets onto open sets?
7. Does an analytic function map closed sets onto closed sets?
8. Let f be an entire function such that $\int_{|z|=R} \frac{f'(z)}{f(z)}\, dz = 0$ for all $R > 200$. Is f a constant?

Exercises 9.59.

1. If $f(z)$ is analytic inside and on the simple closed contour C, and $f(z) \neq a$ on C, show that the number of times $f(z)$ assumes the value a inside C is given by
$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)-a}\,dz.$$
2. Let $f(z)$ be analytic inside and on a simple closed contour C except for a finite number of poles inside C. Denote the zeros by $z_1, \dots, z_n$ (none of which lies on C) and the poles by $w_1, \dots, w_m$. If $g(z)$ is analytic inside and on C, prove that
$$\frac{1}{2\pi i}\int_C g(z)\frac{f'(z)}{f(z)}\,dz = \sum_{j=1}^{n} g(z_j) - \sum_{j=1}^{m} g(w_j),$$
where each zero and pole occurs as often in the sum as is required by its multiplicity.
3. If $P(z) = a_0 + a_1 z + \cdots + a_n z^n$, evaluate
$$\frac{1}{2\pi i}\int_{|z|=R} \frac{zP'(z)}{P(z)}\,dz$$
for large values of R.
4. Using the argument principle, prove the Fundamental Theorem of Algebra.
5. If $f(z)$ is analytic at z_0, show that $f(z)$ has a zero of order k at z_0 if and only if $1/f(z)$ has a pole of order k at z_0.
6. If $f(z)$ is analytic and nonzero in the disk $|z| < 1$, prove that for $0 < r < 1$
$$\exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log|f(re^{i\theta})|\,d\theta\right) = |f(0)|.$$
7. Show that the polynomial $z^4 + 4z - 1$ has one root in the disk $|z| < 1/3$ and the remaining three roots in the annulus $1/3 < |z| < 2$.
8. Find the number of roots of the equation $z^4 - 8z + 10 = 0$ in the unit disk $|z| < 1$ and in the annulus $1 < |z| < 3$, respectively.
9. Show that there exists one root in $|z| < 1$, and three roots in $|z| < 2$ for the equation $z^4 - 6z + 3 = 0$.
10. If $a > e$, show that the equation $e^z = az^n$ has n roots inside the unit circle. When $n = 2$, show that both roots are real.
11. If $a > 1$, prove that $f(z) = z + e^{-z}$ takes the value a at exactly one point in the right half-plane.
12. Show that the equation $z^3 + iz + 1 = 0$ has neither a real root nor a purely imaginary root.
13. Show that the number of roots of the equation $z^4 - 6z + 1 = 0$ in the annulus $1 < |z| < 2$ is 3.

14. Let $F_1(z) = z^5 + z + 16$, $F_2(z) = z^7 - 5z^3 - 12$ and $F_3(z) = z^7 + 6z^3 + 12$. Determine whether all zeros of these functions lie in the annulus $1 < |z| < 2$.
15. Show that the polynomial $z^3 - z^2 + 4z + 5$ has all its roots in the disk $|z| < 3$.
16. How many roots of the equation $z^4 + z^3 + 1 = 0$ have modulus between $3/4$ and $3/2$?
17. Show that, however small R, all the zeros of the function

$$1 + \frac{1}{z} + \frac{1}{2!z^2} + \cdots + \frac{1}{n!z^n}$$

lie in the disk $|z| < R$, if n is sufficiently large.
18. Suppose that $f(z)$ is analytic for $|z| \leq 1$ such that $|f(z)| < 1$ for $|z| = 1$. Show that $f(z) = z^n$ has exactly n solutions in $|z| < 1$.
19. Suppose $\{f_n(z)\}$ is a sequence of analytic functions that converge uniformly to $f(z)$ on all compact subsets of a domain D. Let $f_n(z_n) = 0$ for every n, where each z_n belongs to D. Show that every limit point of $\{z_n\}$ that belongs to D is a zero of $f(z)$.
20. Suppose that $f(z)$ is analytic at z_0 and that $f(z) - f(z_0)$ has a zero of order n at z_0. Show that there exist neighborhoods $N(z_0; \delta)$ and $N(f(z_0); \epsilon)$ such that each point in $N(f(z_0); \epsilon)$ is the image of at least one and at most n distinct points in $N(z_0; \delta)$.
21. Suppose $f(z)$ is analytic at z_0 with $f'(z_0) \neq 0$. Show that there exists an analytic function $g(z)$ such that $f(g(z)) = z$ in some neighborhood of z_0. This is known as the *inverse function theorem*.

References and Further Reading

[A] Ahlfors, L., *Complex Analysis*, 2nd ed., McGraw-Hill, New York, 1966.

[Ap] Apostol, T., *Mathematical Analysis*, Addison-Wesley, Reading, MA, 1967.

[CG] Carlesson, L., Gamelin, T.W., *Complex Dynamics*, Springer, Berlin, Heidelberg, New York, 1993.

[C] Conway, J.B., *Functions of One Complex Variable*, 2nd edition, Springer, Berlin, Heidelberg, New York, 1978.

[DO] Depree, J., Oehring, C., *Elements of Complex Analysis*, Addison-Wesley Publishing Company, Reading, MA, 1969.

[E] Edwards, H., *Riemann Zeta Function*, Academic Press, New York, 1974.

[G] Gamelin, T.W.: *Complex Analysis*, Springer, Berlin, Heidelberg, New York, 2001.

[Ge] Gelbaum, B.R.: *Modern Real and Complex Analysis*, John Wiley & Sons, Inc., 1995.

[GK] Greene, R.E., Krantz, S.G., *Function Theory of One Complex Variable*, John Wiley & Sons, Inc., 1997.

[H] Henrici, P., *Applied and Computational Complex Analysis*, Vol. I, II, III John Wiley & Sons, Inc., 1974, 1977, and 1986.

[Hi] Hille, E., *Analytic Function Theory*, Vol. 2., Ginn, Lexington, MA, 1962.

[I] Ivic, A., *The Zeta Function*, John Wiley & Sons, Inc., 1987.

[LR] Levinson, N., Redheffer, R., *Complex Variables*, Holden-Day, 1970.

[M] Markushevich, A.I., *Theory of Functions of a Complex Variable*, Chelsea, 1977.

[N] Nehari, Z., *Introduction to Complex Analysis*, Allyn and Bacon, Boston, 1968.

[Ne] Newman, M., *Elements of the Topology of Plane Sets of Points*, 4th ed., Cambridge University Press, 1961.

[P] Palka, B.P., *An Introduction to Complex Function Theory*, Springer, Berlin, Heidelberg, New York, 1991.

[P1] Ponnusamy, S., *Foundations of Complex Analysis*, 2nd ed., Narosa Publishing House, India, 2005.

[P2] Ponnusamy, S., *Foundations of Functional Analysis*, Narosa Publishing House, India, 2003.

[R1] Rudin, W., *Principles of Mathematical Analysis*, 2nd ed., McGraw-Hill, New York, 1964.

[R2] Rudin, W., *Real and Complex Analysis*, 3rd ed., McGraw-Hill, New York, 1987.
[S] Sprecher, D., *Elements of Real Analysis*, Academic Press, New York, 1970.
[T] Taylor, A., *Advanced Calculus*, Blaisdell, New York, 1955.
[Ti] Titchmarch, E.C., *Theory of Functions*, Oxford University Press, 1960.
[V] Veech, W.A., *A Second Course in Complex Analysis*, W.A. Benjamin, Inc., New York, 1967.

Index of Special Notations

Symbol	Meaning for
$\emptyset$	empty set
$a \in S$	a is an element of the set S
$a \notin S$	a is not an element of S
$\{x : \dots\}$	set of all elements with the property $\dots$
$X \cup Y$	set of all elements in X or Y; i.e., union of the sets X and Y
$X \cap Y$	set of all elements in X as well as in Y; i.e., intersection of the sets X and Y
$X \subseteq Y$	set X is contained in the set Y; i.e., X is a subset of Y
$X \subset Y$ or $X \subsetneq Y$	$X \subseteq Y$ and $X \neq Y$; i.e., set X is a proper subset of Y
$X \times Y$	Cartesian product of sets X and Y, $\{(x, y) : x \in X,\ y \in Y\}$
$X \setminus Y$ or $X - Y$	set of all elements that live in X but not in Y
X^c	complement of X
$\Longrightarrow$	implies (gives)
$\Longleftrightarrow$	if and only if, or 'iff'
$\longrightarrow$ or $\to$	converges (approaches) to; into
$\not\longrightarrow$ or $\not\to$	does not converge
$\not\Longrightarrow$	does not imply
$\mathbb{N}$	set of all natural numbers, $\{1, 2, \cdots\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\} = \{0, 1, 2, \cdots\}$
$\mathbb{Z}$	set of all integers (positive, negative and zero)
$\mathbb{Q}$	set of all rational numbers, $\{p/q : p,\ q \in \mathbb{Z},\ q \neq 0\}$
$\mathbb{R}$	set of all real numbers, real line
$\mathbb{R}_\infty$	$\mathbb{R} \cup \{-\infty, \infty\}$, extended real line
$\mathbb{C}$	set of all complex numbers, complex plane

$\mathbb{C}_\infty$	extended complex plane, $\mathbb{C} \cup \{\infty\}$
$\mathbb{R}^n$	n-dimensional real Euclidean space, the set of all n-tuples $x = (x_1, x_2, \dots, x_n)$, $x_j \in \mathbb{R}$, $j = 1, 2, \dots, n$
$i\mathbb{R}$	set of all purely imaginary numbers, imaginary axis
$\overline{z}$	$\overline{z} := x - iy$, complex conjugate of $z = x + iy$
$\|z\|$	$\sqrt{x^2 + y^2}$, modulus of $z = x + iy$, x, $y \in \mathbb{R}$
$\mathrm{Re}\, z$	real part x of $z = x + iy$
$\mathrm{Im}\, z$	imaginary part y of $z = x + iy$
$\arg z$	set of real values of θ such that $z = \|z\|e^{i\theta}$
$\mathrm{Arg}\, z$	argument $\theta \in \arg z$ such that $-\pi < \theta \leq \pi$; the principal value of $\arg z$
$\limsup \|z_n\|$	upper limit of the real sequence $\{\|z_n\|\}$
$\liminf \|z_n\|$	lower limit of the real sequence $\{\|z_n\|\}$
$\lim \|z_n\|$	limit of the real sequence $\{\|z_n\|\}$
$\sup S$	least upper bound, or the supremum, of the set $S \subset \mathbb{R}_\infty$
$\inf S$	greatest lower bound, or the infimum, of the set $S \subset \mathbb{R}_\infty$
$\inf_{x \in D} f(x)$	infimum of f in D
$\max S$	the maximum of the set $S \subset \mathbb{R}$; the largest element in S
$\min S$	the minimum of the set $S \subset \mathbb{R}$; the smallest element in S
$f : D \longrightarrow D_1$	f is a function from D into D_1
$f(z)$	the value of the function at z
$f(D)$	set of all values $f(z)$ with $z \in D$; i.e., $w \in f(D) \Longleftrightarrow \exists\, z \in D$ such that $f(z) = w$
$f^{-1}(D)$	$\{z :\ f(z) \in D\}$, the preimage of D w.r.t f
$f^{-1}(w)$	the preimage of one element $\{z\}$
$f \circ g$	composition mapping of f and g
$\mathrm{dist}\,(z, A)$	distance from the point z to the set A i.e., $\inf\{\|z - a\| :\ a \in A\}$
$\mathrm{dist}\,(A, B)$	distance between two sets A and B i.e., $\inf\{\|a - b\| :\ a \in A, b \in B\}$

$[z_1, z_2]$	closed line segment connecting z_1 and z_2; $\{z = (1-t)z_1 + tz_2 :\ 0 \le t \le 1\}$
(z_1, z_2)	open line segment connecting z_1 and z_2; $\{z = (1-t)z_1 + tz_2 :\ 0 < t < 1\}$
$\Delta(a;r)$	open disk $\{z \in \mathbb{C} :\ \lvert z-a\rvert < r\}$ $(a \in \mathbb{C},\ r > 0)$
$\overline{\Delta(a;r)}$	closed disk $\{z \in \mathbb{C} :\ \lvert z-a\rvert \le r\}$ $(a \in \mathbb{C},\ r > 0)$
$\partial\Delta(a;r)$	the circle $\{z \in \mathbb{C} :\ \lvert z-a\rvert = r\}$
Δ_r	$\Delta(0;r)$
Δ	$\Delta(0;1)$, unit disk $\{z \in \mathbb{C} :\ \lvert z\rvert < 1\}$
$\partial\Delta$	unit circle $\{z \in \mathbb{C} :\ \lvert z\rvert = 1\}$
e^z	$\exp(z) = \sum_{n \ge 0} \frac{z^n}{n!}$, an exponential function
$\operatorname{Log} z$	$\ln\lvert z\rvert + i\operatorname{Arg} z,\quad -\pi < \operatorname{Arg} z \le \pi$
$\log z$	$\ln\lvert z\rvert + i\arg z :=\ \operatorname{Log} z + 2k\pi i,\ k \in \mathbb{Z}$
$\dfrac{\partial}{\partial z}$	$\dfrac{1}{2}\left(\dfrac{\partial}{\partial x} - i\dfrac{\partial}{\partial y}\right)$, Cauchy–Riemann operator
$\dfrac{\partial}{\partial \overline{z}}$	$\dfrac{1}{2}\left(\dfrac{\partial}{\partial x} + i\dfrac{\partial}{\partial y}\right)$
f_z	$\dfrac{\partial f}{\partial z}$, partial derivative w.r.t z
$f_{\overline{z}}$	$\dfrac{\partial f}{\partial \overline{z}}$, partial derivative w.r.t z
$\operatorname{Int}(\gamma)$	interior of γ
$\operatorname{Ext}(\gamma)$	exterior of γ
$\gamma_1 + \gamma_2$	sum of two curves γ_1, γ_2
$L(\gamma)$	length of the curve γ
$f^{(n)}(a)$	n-th derivative of f evaluated at a
$\left.\begin{array}{l} f(z) = O(g(z)) \\ \text{as } z \to a \end{array}\right\}$	there exists a constant K such that $\lvert f(z)\rvert \le K\lvert g(z)\rvert$ for all values of z near a
$\left.\begin{array}{l} f(z) = o(g(z)) \\ \text{as } z \to a \end{array}\right\}$	$\displaystyle\lim_{z \to a} \frac{f(z)}{g(z)} = 0$
$\left.\begin{array}{l} \lim\limits_{n\to\infty} z_n = z, \\ \text{or } z_n \to z,\ \text{or} \\ d(z_n, z) \to 0 \end{array}\right\}$	sequence $\{z_n\}$ converges to z with a metric d
$\operatorname{Res}[f(z); a]$	residue of f at a

Index

Hints for Selected Questions and Exercises

Questions 1.1:

1. The set $\{0, 1, 2, \ldots, p\}$, p a prime, is a field under the operations of addition and multiplication modulo p.
2. Between any two elements in an ordered field there is another element.
8. We can see clearly the relationship between a complex number and a point in the plane.
10. Closure. Even though $(1, 1) > (0, 0)$, we have $(1, 1)(1, 1) = (0, 2)$.

Exercises 1.2:

4. (a) $(-5, 14)$ (b) $18-9i$ (c) $-2+2i$ (d) -4 (e) $2^{(n/2)+1} i \sin n\pi/4$.

Questions 1.7:

4. $|z_1+z_2| = |z_1|+|z_2|$ if and only if z_1 and z_2 lie on the same ray emanating from the origin.
7. Because their product is rational.

Exercises 1.8:

1. (b) $13 - 6i$ (d) $\sqrt{2}$ (f) $\sqrt{2}$.
3. (b) $(x + 5)^2 + y^2 > 4^2$ (c) $-1 \leq x \leq 1,\ y = 0$ (d) $y^2 = -20(x - 5)$.
12. We require $|z_1| = |z_2| = |z_1 - z_2|$ so that

$$|z_1|^2 = a^2 + 1 = 1 + b^2 = (a - 1)^2 + (b - 1)^2.$$

This gives $a = \pm b$, $a^2 - 2a + 1 - 2b = 0$, $b^2 - 2b + 1 - 2a = 0$. Note that

$$\begin{aligned} a = \pm b &\Rightarrow b^2 \mp 2b + 1 - 2b = 0 \\ &\Rightarrow b^2 - 4b + 1 = 0 \\ &\Rightarrow b = 2 \pm \sqrt{3} \\ &\Rightarrow b = 2 - \sqrt{3} \quad (\text{ as } 0 < b < 1) \\ &\Rightarrow a = 2 - \sqrt{3} \quad (\text{ as } 0 < a < 1). \end{aligned}$$

Thus the given points form an equilateral triangle if $a = b = 2 - \sqrt{3}$.

13. As $|z_j| = 1$ ($j = 1, 2, 3$), we have $|z_3 - z_1| = |z_2 - z_3|$ iff $z_1 z_3 = z_2 z_3$. The latter equation follows if we multiply $z_3 = -(z_1 + z_2)$ by z_1 and z_2 and compare the resulting equations. Similarly, we can get $|z_3 - z_1| = |z_2 - z_1|$ and the assertion follows.

Questions 1.13:

2. The most important such function is $f(x) = \ln x$.
9. Note that $\arg(1/z) = \arg \overline{z}$.
13. Follow from text. A careful discussion may also be found in DePree and Oehring [DO].
14. Expanding (1.15) gives some useful identities.
18. The roots of unity form a group under multiplication.
20. This is discussed in Chapter 4.

Exercises 1.14:

10. (a) $\pm(1+i)/\sqrt{2}$ (d) $\pm\sqrt[4]{2}(\cos(\pi/8) + i\sin(\pi/8))$ (g) $\pm(1-i)/\sqrt{2}$
11. We have $\omega^{3n+1} = \omega$ and $\omega^{3n+2} = \omega^2$ and so it is easy to see that

$$S_n = \begin{cases} -2\omega & \text{if } n = 3m \\ \dfrac{1-\omega}{1+\omega} & \text{if } n = 3m+1, \\ \dfrac{1}{\omega} & \text{if } n = 3m+2 \end{cases} \quad \text{for } m = 1, 3, 5, 7, \ldots ,$$

and

$$S_n = \begin{cases} 0 & \text{if } n = 3m \\ 1 & \text{if } n = 3m+1, \\ 1-\omega & \text{if } n = 3m+2 \end{cases} \quad \text{for } m = 0, 2, 4, \ldots .$$

12. There is nothing to prove if $\omega = 1$. Therefore, we assume that ω is different from 1. Since ω is a cube root of unity, we have $\omega^2 = -1 - \omega$ and therefore

$$\begin{aligned} &(a + b\omega + c\omega^2)^3 \\ &\quad = (a + b\omega - c(1+\omega))^3 \\ &\quad = (a-c)^3 + 3(a-c)[(b-c)\omega][a - c + \omega(b-c)] + (b-c)^3 \\ &\quad = (a-c)^3 + (b-c)^3 + 3(a-c)(b-c)[(a-c)\omega + (b-c)\omega^2]. \end{aligned}$$

As $(a-c)\omega + (b-c)\omega^2 = (a-c)\omega - (1+\omega)(b-c) = (a-b)\omega - (b-c)$, the right-hand side of the above expression is real iff

$$0 = \operatorname{Im}[(a-b)\omega - (b-c)]$$

and this holds if $a = b$. In this way we see that the required condition is that a, b, c are not all different.

15. As $n \to \infty$ the sum approaches 2π, the circumference of the unit circle.

16. Rewrite the given equation as $((1+z)/(1-z))^5 = 1$, since $z = 1$ is not possible. So, with $z \neq 1$, the solution is given by

$$\frac{1+z}{1-z} = e^{i2k\pi/5}; \quad \text{i.e.,} \quad z = \frac{1-e^{i2k\pi/5}}{1+e^{i2k\pi/5}}, \quad k = 0,1,2,3,4.$$

Writing the above equation as

$$z = \frac{e^{-ik\pi/5} - e^{ik\pi/5}}{e^{-ik\pi/5} + e^{ik\pi/5}} = \frac{-2i\sin(k\pi/5)}{2\cos(k\pi/5)}, \quad k = 0,1,2,3,4,$$

we see that all the roots of the given equation lie in the imaginary axis at $z_k = -i\tan(k\pi/5)$, $k = 0,1,2,3,4$.

17. The roots of equation $(z-1)^5 = -1$ are the vertices of a regular pentagon having center at 1 and vertex at the origin respectively. Comparing the above equation with the given equation we obtain $\alpha = -5$, $\beta = 10$, $\gamma = -10$, $\delta = 5$ and $\eta = 0$.

18. Since $|(1+ix)/(ix-1)| = 1$, the equation $((ix+1)/(ix-1))^n = \zeta$ $(\zeta = e^{i\theta})$ becomes

$$\frac{ix+1}{ix-1} = e^{i(\theta+2k\pi)/n},$$

that is

$$ix = \frac{1+e^{i(\theta+2k\pi)/n}}{-1+e^{i(\theta+2k\pi)/n}} = \frac{2\cos[(\theta+2k\pi)/2n]}{2i\sin[(\theta+2k\pi)/2n]},$$

where $k = 0, 1, \ldots, n-1$. So

$$x = \cot\left(-\left(\frac{\theta+2k\pi}{2n}\right)\right), \quad \text{i.e.,} \quad \theta = -2k\pi - 2n\cot^{-1}(x).$$

Since $\zeta = e^{i\theta}$, we have

$$\left(\frac{ix-1}{ix-1}\right)^n = e^{i\theta} = e^{i(-2k\pi - 2n\cot^{-1} x)} = e^{-2in\cot^{-1}(x)}$$

proving the assertion.

Questions 2.9:

4. Any set of real numbers that is closed is also a closed subset of the plane. The empty set is the only set that is open in both the real line and the plane.
6. The integers have no limit point.
7. A boundary point of a connected set with more than one point must be a limit point.
8. $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$. To see that the containment is proper, let A denote the set of irrational numbers and B denote the set of rational numbers.
10. This is known as a convex set.

11. This is known as a starlike set. The plane minus the negative real axis is starlike with respect to the origin, but is not convex.

Exercises 2.10:

4. (a) Bounded open set.
   (b) Not open, not closed, not connected, not bounded.
   (e) Open on the real line, connected, unbounded.

Questions 2.22:

2. Let $x_n = (-1)^n$ and $y_n = (-1)^{n+1}$.
4. The sequence $\{x_n\}$, where

$$x_n = \begin{cases} 1/n & \text{if } n \text{ is odd} \\ n & \text{if } n \text{ is even} \end{cases},$$

   is an unbounded sequence with a limit point.
5. When the set is closed.
6. $2, 0, 1, 1, 1, \ldots$.
7. The set of rational numbers may be expressed as a sequence.
10. The sequence $\{b_n\}$ is increasing.

Exercises 2.23:

2. Construct disjoint neighborhoods about two distinct limit points.
4. (b), (c), (d).
5. Suppose that $z_n \to z_0$. Then, given $\epsilon > 0$ there exists an N (assume $N \geq 2$) such that $|z_n - z_0| < \epsilon/2$ for all $n \geq N$. Now for each $n \geq N$, we have

$$\begin{aligned} \left| \frac{1}{n} \sum_{k=1}^{n} z_k - z_0 \right| &= \left| \frac{1}{n} \sum_{k=1}^{n} (z_k - z_0) \right| \leq \frac{1}{n} \sum_{k=1}^{n} |z_k - z_0| \\ &= \frac{1}{n} \sum_{k=1}^{N-1} |z_k - z_0| + \frac{1}{n} \sum_{k=N}^{n} |z_k - z_0| \\ &< \frac{1}{n} \sum_{k=1}^{N-1} |z_k - z_0| + \frac{1}{n}(n - (N-1)) \frac{\epsilon}{2} \\ &< \epsilon, \qquad \text{whenever } n \geq \max\left\{ N, \frac{2}{\epsilon} \sum_{k=1}^{N-1} |z_k - z_0| \right\}. \end{aligned}$$

6. (a) $1, 2, \frac{1}{2}, 2, \frac{1}{3}, 2, \ldots$
   (b) $1, 2, \ldots, n, \frac{3}{2}, \frac{5}{2}, \ldots, \frac{2n+1}{2}, \ldots, \frac{k+1}{k}, \frac{2k+1}{k}, \ldots, \frac{nk+1}{k}, \ldots$
   (c) A sequence consisting of all the rational numbers.
8. $1, 2, 2\frac{1}{2}, 3, 3\frac{1}{3}, 3\frac{2}{3}, 4, 4\frac{1}{4}, 4\frac{2}{4}, 4\frac{3}{4}, 5, 5\frac{1}{5}, \ldots$.

Questions 2.28:

2. The infinite union may not be compact. Each integer is compact, but their union is not. Also, $[0, 1-1/n]$ is compact for each n, but $\bigcup_{n=2}^{\infty}[0, 1-1/n]$ is not compact.
3. It is open and unbounded.
5. If there are no limit points.
6. Infinitely many.

Exercises 2.29:

6. For p and q relatively prime positive integers, $f(p, q) = 2^p 3^q$ is a one-to-one mapping of the positive rationals into the positive integers.
7. For each open set in the cover, choose a point in the set both of whose coordinates are rational.

Questions 2.30:

2. A set containing a neighborhood of ∞ is unbounded, but the converse does not hold. For instance, $\{z : \operatorname{Re} z > 0\}$ does not contain a neighborhood of ∞.
6. They are identified with themselves.
8. The $n + 1$ sphere.
11. The image of the line $x + y = 1$ is given by the plane $x_1 + y_1 + u_1 = 1$, where $x_1^2+y_1^2+u_1^2 = 1$. This is the intersection of the plane $x_1+y_1+u_1 = 1$ and the Riemann sphere, which is a circle passing through the north pole $(0, 0, 1)$.

Exercises 2.31:

1. If z_0 is a limit point, then $N(\infty; |z_0| + 1)$ does not contain infinitely many points of the sequence.
4. $\left(\dfrac{x_1}{1-u_1}, \dfrac{y_1}{1-u_1}\right)$.
6. Review Exercises 5 and 6 after reading Chapter 3.
7. See the book by S. Ponnusamy [P1].

Questions 2.45:

1. It might be confused with our definition of a domain as an open connected set.
4. If points are always closer in the w plane, the function is uniformly continuous. The converse is not true. Consider $f(z) = 2z$ on a bounded set.
5. All sequences are uniformly continuous. Just choose $\delta = 1$.
7. It need not be a limit point; for instance, a constant function.
9. No such thing. The function is uniformly continuous on $|z| \geq \epsilon$ for all $\epsilon > 0$.
10. A mapping from the unit disk onto two points.

Exercises 2.46:

1. (a) $6i$ (b) $-(8+6i)/5$ (c) 0 (d) 1/2.
2. (a) is continuous whereas (b) is uniformly continuous; (c) and (d) are discontinuous at the origin.
8. $f(z) = \sin \pi z$.
18. Use Exercise 12.

Questions 3.2:

3. A half-plane.
5. $\dfrac{1}{z+b} \neq \dfrac{1}{z} + b$.
6. Rotation and magnification.

Exercises 3.3:

1. (a) $v = -3(u-2)$ (d) $(u-3)^2 + (v-1)^2 = 8$.
2. (a) $\mathrm{Im\,w} > -1$ (b) $\mathrm{Im\,w} > 0$.
3. The strip between the lines $v = u - 3$ and $v = u - 7$.
6. (c) The triangle with vertices $-1 + 11i$, $-13 + 5i$, and $2 - 10i$.
7. (b) $(u - \frac{1}{2})^2 + (v+1)^2 = \frac{5}{4}$.
8. (a) $(u - \frac{2}{3})^2 + v^2 = (\frac{1}{3})^2$
   (d) $(u - \frac{1}{2})^2 + v^2 < \frac{1}{4}$.
9. $0 \le \mathrm{Re}\, z \le 2$ maps onto the right half-plane minus the disk $|w - 1/4| < 1/4$.

Questions 3.27:

3. A linear transformation.
13. Use the fact that lines and circles map onto lines and circles. Review this question after reading Chapter 11.
15. $f(z) = \begin{cases} 0 & \text{for } z = \infty \\ \infty & \text{for } z = 0 \\ z & \text{otherwise.} \end{cases}$

    If we require continuity, the function must be bilinear.

Exercises 3.28:

2. (a) $w = \dfrac{iz-1}{-3z+i}$ (d) $w = -\dfrac{z+2(1-i)}{z-2}$.
4. (a) $\mathrm{Im}\, v \le \frac{1}{2}$ (b) $u^2 + (v - \frac{1}{2})^2 \ge (\frac{1}{2})^2$.
5. (c) $(u-1)^2 + (v-1)^2 < 1$ (d) $u^2 + (v+1)^2 > 2$.
9. What is the image of the real line under a bilinear map?
11. Choose $z_2 = \infty$ in the previous exercise.
12. Suppose $w = (az+b)/(cz+d)$. If $a = d$, $b = c = 0$, then there are infinitely many fixed points. If $(a-d)^2 = -4bc$, then there is one fixed point.

18. For instance, set

$$T(z) = z+1,\ S(z) = z/(z+1) \quad \text{and} \quad U(z) = (z+1)/z.$$

Then

$$T(S(z)) = \frac{2z+1}{z+1},\ S(T(z)) = \frac{z+1}{z+2}$$

and

$$T(U(z)) = \frac{2z+1}{z},\ U(T(z)) = \frac{z+2}{z+1}.$$

19. $w = A(iz - z_0)/(iz - \overline{z}_0)$, $|A| = 1$.
24. We present a direct proof. Clearly the equation of the line L is $y = x+1$ and the equation of the line passing through $3i$ and $2+i$ is $y = -x+3$. Note that these two lines are perpendicular. Solving these two equations give $1+2i$ as its point of intersection. Note that

$$|(1+2i) - 3i| = |1+2i-(2+i)| = \sqrt{2}.$$

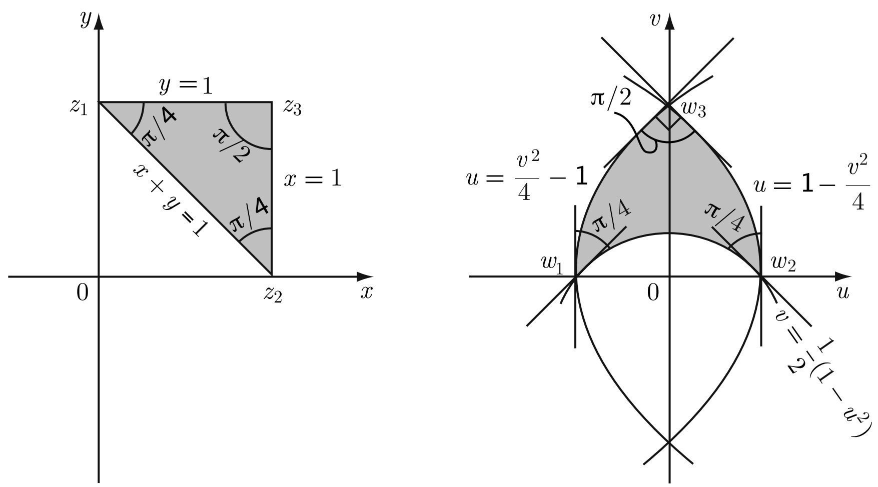
Thus z_1 and z_2 are inverses with respect to the given line.
26. $(3+6i)/5$.

Questions 3.29:

1. It doesn't.
2. Any half-plane whose boundary passes through the origin.
4. When the ray (extended) passes through the origin.

Exercises 3.30:

2. We have $f(z) = x^2 - y^2 + 2ixy = u+iv$, where $u = x^2 - y^2$ and $v = 2xy$. Note that



$$x = 1 \Rightarrow u = 1 - y^2,\ v = 2y, \quad \text{i.e.,} \quad u = 1 - (v^2/4)$$
$$y = 1 \Rightarrow u = x^2 - 1,\ v = 2x, \quad \text{i.e.,} \quad u = (v^2/4) - 1$$
$$x + y = 1 \Rightarrow \begin{cases} u = x^2 - (1-x)^2 = 2x - 1 \\ v = 2x(1-x) \end{cases}, \quad \text{i.e.,} \quad v = (1-u^2)/2.$$

4. Rewrite the given function as
$$\frac{z}{(1-z)^2} = \frac{1}{4}\left(\frac{1+z}{1-z}\right)^2 - \frac{1}{4}.$$
7. Set $w_1 = T_1(z) = z^n$ and $w_2 = T_2(z) = e^{i\alpha}(z - z_0)/(z - \overline{z}_0)$, where $\alpha \in \mathbb{R}$ and z_0 with $\operatorname{Im} z_0 > 0$ are fixed (choose for example, $\alpha = 0$ and $z = i$). Then the composed mapping $w = (T_2 \circ T_1)(z)$ gives a mapping with the desired property.
8. (a) The upper half-plane. (b) The plane minus the positive real axis.
9. The unit disk n times.

Questions 4.7:

2. Infinite strips of width 2π.
4. Unbounded along any ray other than one along the real axis.
8. No, as will be shown in Chapter 8.
11. e^z is unbounded along every ray in the right half-plane, while $e^z + z$ is unbounded along every ray.
14. $\tan z = i \iff e^{iz} - e^{-iz} = i^2(e^{iz} + e^{-iz}) \iff e^{iz} = 0$.

Exercises 4.8:

1. (a) $2k\pi/3i$ (b) $\pm(1+i)\sqrt{k\pi}$ $(k \geq 0)$
   (c) $\ln|2k\pi| + i(\pi/2 + n\pi)$.
6. (a) $e^{x/(x^2+y^2)}\left(\cos\dfrac{y}{x^2+y^2} - i\sin\dfrac{y}{x^2+y^2}\right)$ (b) $|e^{1/z}| \leq e^{1/\epsilon}$.
8. (c) $|\sin z|^2 + |\cos z|^2 \geq |\sin^2 z + \cos^2 z| = 1$.

Questions 4.10:

1. The inverse of the exponential function.
2. Yes.
4. The further in the right half-plane, the larger is the area of its image.
9. $\exp(f(z))$.
11. At all points except $z = \pi/2 + 2k\pi$. This will be better understood after Chapter 10.

Exercises 4.11:

1. (a) The line segment from $[(1+i)/\sqrt{2}]e^{-5}$ to $[(1+i)/\sqrt{2}]e^{5}$.
   (c) The part of the annulus in the upper half-plane bounded by the semicircles $|w| = 1/e^2$ and $|w| = e$.

2. (a) The region $|w| \leq 1$, $\operatorname{Im} w \geq 0$.
   (b) The region $|w| \leq 1$, $\operatorname{Re} w \leq 0$.
   (c) The region $|w| \geq 1$, $\operatorname{Re} w \geq 0$.

Questions 4.14:

1. The imaginary part of the logarithm is the argument.
2. This is perfectly consistent, although sometimes inconvenient.
7. Only if $-\pi < \operatorname{Arg} z_1 - \operatorname{Arg} z_2 \leq \pi$.

Exercises 4.15:

1. (c) $\frac{1}{2}\ln(x^2 + y^2) + i\tan^{-1}(y/x)$.

Questions 4.22:

1. No, because $2k\pi \neq 0$.
3. $(a + bi)^{c+di}$ is real if $d \ln|a + bi| + c \arg(b/a) = k\pi$, k an integer.
4. It assumes at most mn distinct values.
6. Because the function is not single-valued in any neighborhood of the origin. The function is also discontinuous at the origin.
8. One is an n-valued function, and the other is an n-to-one mapping.
9. Only when m and n are relatively prime. For instance, $(z^2)^{1/2}$ has two vales, whereas $(z^{1/2})^2$ has only one.

Exercises 4.23:

1. (b) $\pi^e e^{i(\pi/2+2k\pi)e}$ (c) $\frac{1}{2}$.
8. (c) $\frac{1}{2}\tan^{-1}(2/-1) + (i/4)\ln 5$.

Questions 5.13:

2. Because only one “bad” path need be found.
7. The derivative of

$$f(z) = \begin{cases} x^2 \sin 1/x & \text{if } x \neq 0 \\ 0 & \text{if } x = 0, \end{cases}$$

   exists but is not continuous at the origin.
8. Usually when we are involved with expressions like $x^2 + y^2$.
11. If $f'(z)$ exists, then $f(z)$ is constant.
12. Nowhere when $f(z) = |z|$. Nowhere except at $z = 0$ when $f(z) = |z|^2$.

Exercises 5.14:

1. (c) and (d).
2. (a), (b), (c), (e), (f) differentiable at the origin, (d) differentiable everywhere.
7. (a) $\dfrac{(x^2 + y^2)^{n/2}}{y}$ (b) $y/\sqrt{x}$.

Questions 5.28:

2. Local versus global behavior.
4. A function that is differentiable everywhere in the plane.
11. If $f'(z)/g'(z)$ is continuous at z_0.
12. No. $1/z$, $0 < |z| < 1$.
14. Because the argument is not defined.
20. $(\operatorname{Arg} z)^2$ is continuous on $\mathbb{C}\setminus\{0\}$ and $(\operatorname{Arg} z)^3$ is discontinuous on the negative real axis.
21. No.

Exercises 5.29:

2. (a) $a = b$, $c = -1$ (b) $a = b = c/2$ (c) $a = 1$, $b = 2k\pi$ (d) $a = b = 0$.
8. Let $f = u + iv$ be entire. Then (see also Example 5.25), as
$$f'(z) = u_x + iv_x \equiv v_y - iu_y,$$
it follows that $u_x = 0 = v_y$ and so, $u = \phi(y)$ and $v = \psi(x)$. But then,
$$f(z) = \phi(y) + i\psi(x) \quad \text{and} \quad f'(z) = i\psi'(x) = -i\phi'(y)$$
which shows that $\operatorname{Re} f'(z) = 0$ and therefore, $f'(z)$ is a constant. Hence, $f(z) = az + b$ for some constants a and b with $\operatorname{Re} a = 0$.
16. (a) $1/3$ (b) 0 (c) 2 (d) Does not exist.

Questions 5.40:

1. If a property holds for analytic functions whenever it holds for its real and imaginary parts.
4. If $f(z)$ is analytic, then $|f(z)|$ is continuous and $\ln|f(z)|$ is harmonic when $f(z) \neq 0$.
5. No. See Chapter 10 for details.

Exercises 5.41:

2. (a) $v = ay - bx + c$ (b) $v = \dfrac{x}{x^2 + y^2} + c$ (c) $v = 3x^2y - y^3 + c$ (d) $v = -\ln|z| + c$ (e) $v = e^{x^2 - y^2} \sin 2xy + c$.
4. Follow the idea of Example 5.35.
6. We have $v(x, y) = (y^2 - x^2)/2 + k$ and $u + iv = -iz^2/2 + ik$, where k is some real constant.
7. $a = 3$, $v = 3xy^2 + \dfrac{y^2}{2} - x^3 - \dfrac{x^2}{2} + c$.
9. Note that $u_x = 3ax^2 + y^2 + 1$, $u_y = 2xy$ and so
$$u_{xx} + u_{yy} = 6ax + 2x = 2x(3a + 1) = 0$$
gives $a = -1/3$ to make u harmonic in $\mathbb{C}$. As a derivative formula for $f = u + iv$ is given by $f'(z) = u_x(x, y) - iu_y(x, y)$, it follows that $f'(z) = -x^2 + y^2 + 1 - 2ixy = -z^2 + 1$. This gives

$$f(z) = -(z^3/3) + z + c$$

where c is a constant. Note that harmonic conjugates v are given by

$$v(x, y) = \operatorname{Im}(-z^3/3 + z + c) = -(1/3)(3x^2y - y^3) + y + \operatorname{Im} c.$$

13. $u_x v_x = -u_y v_y$.
14. As $\operatorname{Im}\left(f^2(z)\right) = 2uv$ and $f^2(z)$ is analytic on D, uv is harmonic.
19. (a) $-\dfrac{1}{\sqrt{1-z^2}}$ (b) $\dfrac{1}{1+z^2}$ (c) $\dfrac{1}{\sqrt{z^2(z^2-1)}}$.

Questions 6.18:

5. If $a_n > 0$ and $\sum_{n=1}^{\infty} a_n$ converges, then there exists a positive sequence $\{b_n\}$ such that $\sum_{n=1}^{\infty} b_n$ converges with $a_n/b_n \to 0$.
8. See Exercises 6.19 (8) and 6.19 (9).
10. All sequences have a limit superior, but we have to avoid expressions like $\infty - \infty$.
12. The conclusion is valid as long as one sequence does not approach ∞ while the other approaches $-\infty$.
13. No. $(1/2n)^{1/n} < 1$ for every n.

Exercises 6.19:

4. (a) Set $a_n = r_n - r_{n+1}$, and apply the Cauchy criterion.
   (b) Show that $a_n/\sqrt{r_n} < 2(\sqrt{r_n} - \sqrt{r_{n+1}})$.

Questions 6.36:

2. The sequence $\{z + 1/n\}$ is unbounded in the plane and converges uniformly.
5. A point is a compact set.
7. Define $f_n(z) = \begin{cases} 1/n & \text{if } z \text{ is real,} \\ 0 & \text{otherwise.} \end{cases}$
   The sequence $\{f_n\}$ converges uniformly to zero in the plane.
8. $f_n(z) = (-1)^n$ does not converge, although $|f_n(z) - 1| = 0$ for infinitely many n.
10. $f_n(z) = \sum_{k=1}^{n}(z^k/k^2)$ converges uniformly in $|z| \leq 1$, but the limit function is not differentiable at $z = 1$.

Exercises 6.37:

8. The sequence $\{z/n\}$ converges uniformly to 0 on $|z| \leq 1$, but $\sum_{n=1}^{\infty}(z/n)$ diverges for $z \neq 0$.
10. (a) Converges uniformly to 0, where defined.
    (b) Converges uniformly for $\operatorname{Re} z \geq \epsilon > 0$ and pointwise for $\operatorname{Re} z > 0$.
    (c) Converges uniformly for $\operatorname{Re} z \leq 0$.
    (d) Converges uniformly to 1 for $|z| \leq r < 1$ and pointwise to 1 for $|z| < 1$; converges uniformly to 0 for $|z| \geq R > 1$ and pointwise to 0 for $|z| > 1$; converges to $\frac{1}{2}$ when $z = 1$.

11. (a) Absolutely for $|z| < 1$ and uniformly for $|z| \leq r < 1$.
    (b) Absolutely for $|z^2 + 1| > 1$ and uniformly for $|z^2 + 1| \geq R > 1$.
    (c) Absolutely where defined ($z \neq n$), and uniformly on bounded subsets of the plane that exclude disks centered at the integers.
    (d) Absolutely for $|z| > 1$ and uniformly for $|z| \geq R > 1$.

Questions 6.57:

1. $|z_0| \leq |z_1|$.
2. An entire function.
6. If $\sum_{n=0}^{\infty} na_n z^n$ converges, then $\sum_{n=0}^{\infty} a_n z^n$ converges. The converse is false.
7. It is at least 1.
9. $f_n(z) = z + n$ does not converge, although $f_n'(z) = 1$ converges.
11. In Chapter 8, it will be shown that all analytic functions have power series representations.

Exercises 6.58:

3. $|a_n z^n| \leq |a_0|\,|z|^n$ for all n.
5. Either $\sum_{n=0}^{\infty} a_n$ diverges or the series is a polynomial.
10. (a) $|a|$ (b) $1/|a|$, $|a| > 1$; 1, $|a| \geq 1$ (c) 1 (d) $1/e$ (e) 1; For (f), because of the presence of factorials, it is more convenient to use ratio test. Now

$$\frac{1}{R} = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} \frac{(n+1)^2}{(2n+2)(2n+1)} = \frac{1}{4}.$$

11. (a) 2 (b) $1/\sqrt{3}$ (c) ∞ (d) 1 (e) 5/3; For (f), we note that

$$|2^n z^{n^2}|^{1/n} = 2|z|^n \to \begin{cases} 0 \text{ if } |z| < 1 \\ 2 \text{ if } |z| = 1 \\ \infty \text{ if } |z| > 1 \end{cases}$$

    and therefore, the series converges for $|z| < 1$.
12. Let $s_n = \sum_{i=1}^{n} a_i$, and consider $\sum_{i=m}^{n} a_i b_i = \sum_{i=m}^{n} (s_i - s_{i-1}) b_i$.
14. (b) $9 - 2(z+2) - 3(z+2)^2 + (z+2)^3$.
16. Given $R^{-1} = \limsup_{n\to\infty} |a_n|^{1/n} = \limsup_{n\to\infty} |1/a_n|^{1/n}$. It is easy to see that

$$\limsup_{n\to\infty} \frac{1}{|a_n|^{1/n}} = \frac{1}{\limsup |a_n|^{1/n}} = R$$

    and so the last equation gives $R^2 = 1$.
17. Let $f(z) = \sum_{n=0}^{\infty} \cos(n\pi/3) z^n$. We first compute

$$a_n = \cos(n\pi/3) = \begin{cases} (-1)^k & \text{if } n = 3k \\ (-1)^k/2 & \text{if } n = 3k+1\,, \\ (-1)^{k+1}/2 & \text{if } n = 3k+2 \end{cases} \quad k \in \mathbb{N}_0,$$

so that

$$\begin{aligned} f(z) &= 1 + \left(\frac{z}{2} - \frac{z^2}{2} - z^3\right) - \left(\frac{z^4}{2} - \frac{z^5}{2} - z^6\right) + \cdots \\ &= 1 + \left(\frac{z}{2} - \frac{z^2}{2} - z^3\right)(1 - z^3 + z^6 - \cdots) \\ &= 1 + \left(\frac{z}{2} - \frac{z^2}{2} - z^3\right)\frac{1}{1+z^3} \\ &= \frac{(1+z)(1-z/2)}{1+z^3} \\ &= \frac{1-z/2}{1-z+z^2}. \end{aligned}$$

Questions 6.66:

2. At all points where the denominator is nonzero. This follows from the fact (proved in Chapter 8) that an analytic function admits a power series expansion.
4. $1/(1-z)$ is analytic for $z = 2$. A function cannot be analytic everywhere on $|z| = R$, which is proved in Chapter 13.
8. The radius of convergence of the Taylor series about a point is the distance between that point and the nearest zero of the denominator.

Exercises 6.67:

1. (a) Inequality holds when
   $a_n = \begin{cases} 1 & \text{if } n \text{ is odd} \\ 1/2^n & \text{if } n \text{ is even} \end{cases}$, and $b_n = \begin{cases} 1/2^n & \text{if } n \text{ is odd} \\ 1 & \text{if } n \text{ is even} \end{cases}$.
2. (a) R (b) R (c) ∞ (d) 0.
4. If $a_n \equiv 1$, then $R = 1$ for both series. If $a_n = 2^n$, then $R = 2^{-1/k}$ for $\sum_{n=0}^{\infty} a_n z^{kn}$ and $R = 1$ for $\sum_{n=0}^{\infty} a_n z^{n^2}$.
5. (b) $\limsup\limits_{n\to\infty} \left|n^k/n!\right|^{1/n} = 0$.
6. (a) $\frac{5}{2}$ (b) $\frac{13}{12}$ (c) $5 - 5i$.
7. Note that radius of convergence of $\sum_{n=1}^{\infty} \frac{z^n}{n^2}$ and $\sum_{n=1}^{\infty}(-3)^n z^n$ are 1 and 1/3, respectively. According to Theorem 6.62, the radius of convergence R of the sum of these two series must be at least 1/3. Is $R = 1/3$?
11. Substituting $z = 0$ in the functional equation gives

$$f(0)[1 - f(0)] = 0, \quad \text{i.e., either } f(0) = 0 \text{ or } f(0) = 1.$$

If $f(0) = 0$, then, by differentiating the functional equation, we find that $f'(2z) = f(z)f'(z)$ which gives $f'(0) = 0$. Continuing this process, we get $f^{(k)}(0) = 0$ for all $k \in \mathbb{N}$. Thus, $f(z) = 0$ on Δ_r. If $f(0) = 1$, then by differentiating the last equation we have

$$2f''(2z) = (f'(z))^2 + f(z)f''(z)$$

so that $f''(0) = (f'(0))^2$. In this way, we conclude that

$$f(z) = \sum_{n=0}^{\infty} \frac{(f'(0))^n}{n!} z^n = \exp(f'(0)z).$$

Questions 7.12:

2. No, because the initial and terminal points coincide.
4. The complement of a simply connected domain is connected.
7. The plane minus the integers.
11. It is the continuous image of a compact and connected set.

Exercises 7.13:

1. A circle if $a = b$ and an ellipse if $a \neq b$.
3. As $R \to \infty$, $z(t)$ becomes a circle centered at the origin with radius 1.
5. (a) $z(t) = t + i(1 - 2t) \quad (0 \leq t \leq 1)$
   (d) $z(t) = 1 + 2\cos t + i2\sin t \quad (-2\pi/3 \leq t \leq 2\pi/3)$.
6. $z(t) = t + i(2t^2 - 3) \quad (-1 \leq t \leq 2)$.
7. (b) $4r^2 \cos^2\theta + r^2 \sin^2\theta = 1$.
8. (a) $2\pi i$ (b) $4\pi i$ (e) $-30 + 25\pi i$.
9. (b) $2\pi i$ (d) $2\pi + 4\pi i$.

Questions 7.28:

3. A finite number of discontinuities will not prove significant.
4. Yes, because a contour is compact.
5. As f is continuous at the origin, given $\epsilon > 0$ there exists a $\delta' > 0$ such that $|f(\delta e^{i\theta}) - f(0)| = |f(\delta e^{i\theta})| < \epsilon$ for $\delta < \min\{\delta', r\}$, $\theta \in [0, 2\pi]$. So,

$$\left| \int_0^{2\pi} f(\delta e^{i\theta})\, d\theta \right| \leq \int_0^{2\pi} \left| f(\delta e^{i\theta}) \right| d\theta < 2\pi\epsilon \to 0 \quad \text{as} \quad \epsilon \to 0.$$

   Similarly,

$$\left| \int_{|z|=\delta} \frac{f(z)}{z}\, dz \right| = \left| i \int_0^{2\pi} f(\delta e^{i\theta})\, d\theta \right| \to 0 \quad \text{as} \quad \delta \to 0.$$

7. The parametrization is easier to deal with.

Exercises 7.29:

4. (a) 12π (b) $\sqrt{2}(e^{\pi} - e^{-\pi})$.
5. (b) $\displaystyle\int_C x\, dz = \frac{1-i}{2}, \quad \int_C y\, dz = -\frac{1-i}{2}, \quad \int_C \overline{z}\, dz = 1,$
   (c) $\displaystyle\int_C x\, dz = \pi i, \quad \int_C y\, dz = -\pi, \quad \int_C \overline{z}\, dz = 2\pi i.$

6. Along the line segment from the origin to $1+i$,

$$\int_C z\,dz = i,\quad \int_C |z|\,dz = \frac{1+i}{\sqrt{2}},\quad \int_C z|dz| = \frac{1-i}{\sqrt{2}},\quad \int_C |z|\,|dz| = 1;$$
$$\int_{|z|=1} z\,dz = 0 = \int_{|z|=1} |z|\,dz = \int_{|z|=1} z|dz|,\quad \int_{|z|=1} |z|\,|dz| = 2\pi.$$

9. $2\pi i$.
10. (a) $e^{2+2i}-1$ (b) $e^{1-i}-e^{-(1-i)}+2(1-i)$ (c) $\frac{i}{2}(e^{1-i}-e^{-1+i})$ (e) $(-7+5i)/3$.
11. $\frac{4}{3}$.
12. We have

$$\int_C z|z|\,dz = \int_{-R}^{R} x|x|\,dx + \int_0^{\pi} (Re^{i\theta})(R)iRe^{i\theta}\,d\theta = iR^3\int_0^{\pi} e^{2i\theta}\,d\theta = 0.$$

Questions 7.37:

2. In the use of the Fundamental Theorem of Calculus.
5. See Question 7.37 (2).
7. It need not be analytic: $\int_{|z|=1}(1/z^2)\,dz = 0$.
9. See Section 9.3.

Exercises 7.38:

1. (a) $16/3$ (b) Traversed in the positive sense $-5/3$. (c) 0 (d) $\pi r^2/4$ (e) $128/5$.
6. Regardless of the contour chosen: (a) 0 (b) $i(1+1/e)$.

Questions 7.55:

1. Not necessarily, because $|f(z)|$ is not analytic.
3. $\displaystyle\int_{|z|=1}\frac{dz}{z} \neq 0$.
7. In order to apply the Cauchy-Riemann equations.
10. Any integral multiple of $2\pi i$.
11. If $g_0(z)$ is a solution, then so is $g_0(z)+2k\pi i$.

Exercises 7.56:

3. $f(z) = 1/z^2$.
5. (a) π (b) $-\pi$ (c) 0 (d) 0.
6. (a) $2\pi i$ (b) 0 (c) $2\pi i$. These solutions will be easy to verify after reading Section 8.1.
9. As $\operatorname{Re} z = (z+1/z)/2$ for $|z|=1$, we have

$$I = \int_{|z|=1} (z^2+1+2\alpha z)\frac{f(z)}{z^2}\,dz = 2\pi i(f'(0)+2\alpha f(0)).$$

10. As $f(z)=a^z=e^{z\operatorname{Log} a}$ is an entire function, its primitive is given by

$$F(z)=\frac{e^{z\operatorname{Log} a}}{\operatorname{Log} a}=\frac{a^z}{\operatorname{Log} a}.$$

Questions 8.23:

4. If they are analytic at the point, they are identical.
6. $\{z^n\}$ does not converge uniformly on $|z|<1$.
8. They need not be analytic.
10. $\int_{|z|=1}|z|^2\,dz=0$, but $|z|^2$ is not analytic. Morera's theorem is not applicable because the integral is not zero along every contour.
12. $f(z)=\sum_{n=0}^{\infty} z^n=\dfrac{1}{1-z}\quad (|z|<1)$, $f'(2)=1\neq \sum_{n=1}^{\infty} nz^{n-1}$. Here, $1/(1-z)$ is not defined by $\sum_{n=0}^{\infty} z^n$ at $z=2$. See Chapter 13.

Exercises 8.24:

3. (a) $2\pi i e^2$ (b) $2\pi i e^4$ (c) $8\pi i e^4$ (d) $2\pi i e^2(\sin 2+\cos 2)$ (e) $2\pi e^{-2}\sin 2$ (f) $144\pi i$.
4. (a) 0 (b) 0 (c) 0 (d) $12\pi i$.
6. As $|z|=1$, for the first integral, we may rewrite $\operatorname{Re} z=(z+1/z)/2$. For the second and third integrals, we write $z-1=e^{i\theta}$ so that

$$\overline{z}=1+e^{-i\theta}=1+\frac{1}{z-1}=\frac{z}{z-1}$$

and

$$\operatorname{Im} z=\operatorname{Im}(z-1)=\frac{z-\overline{z}}{2i}=\frac{1}{2i}\left(\frac{z^2-2z}{z-1}\right).$$

Now, use the Cauchy integral formula.

7. (a) $\dfrac{2\pi}{17}$ (c) $\dfrac{\pi}{17}-\dfrac{4\pi i}{17}$ (e) $\dfrac{2\pi}{17}-\dfrac{8\pi i}{17}$
8. Use the Cauchy theorem for multiply connected domains and the Cauchy integral formula.
9. (a) $z+z^2+\frac{1}{3}z^3-\frac{1}{30}z^5$
   (c) $1+z+\frac{3}{2}z^2+\frac{7}{6}z^3+\frac{25}{24}z^4+\frac{27}{40}z^5$.
11. (a) $\frac{1}{2}$ (b) 1 (c) 1 (d) 0.
16. Express as $e^{\alpha\log(1-z)}$ and expand.

Questions 8.55:

2. No. Indeed if $f(z)=\sum_{n=0}^{\infty} a_n z^n$ with $|a_n|\geq n!$, then

$$\frac{1}{R}=\limsup_{n\to\infty}|a_n|^{1/n}\geq \limsup_{n\to\infty}(n!)^{1/n}=\infty,\quad \text{i.e.,}\quad R=0.$$

5. e^z is bounded for $\operatorname{Re} z\leq 0$.
6. No, as will be shown in Chapter 11.

12. If $f(z) \neq 0$ in $\mathbb{C}$, then $\phi(z) = 1/f(z)$ is entire and $|\phi(z)| \to 0$ as $|z| \to \infty$. So, ϕ is entire and bounded on $\mathbb{C}$. Consequently, ϕ and hence f is constant, a contradiction. Thus, f has a zero in $\mathbb{C}$.
13. No, because $\sin \pi z$ and 0 agree at the integers.
17. Consider $\phi(z) = e^{-(a-ib)f(z)}$. Then $|\phi(z)| = e^{-(au+bv)} \leq e^{-c} = M$.

Exercises 8.56:

11. Set $h(z) = f(z)/g(z)$ and show that $h'(z) = 0$.
12. For instance, $f(z) = \sin(1/(z-1))$, $\sin((1+z)/(1-z))$.
15. Define $g(z) = f(z) - \overline{f(\overline{z})}$. Then g is entire and $g(a_n) = 0$ for all $n \geq 1$. As every bounded sequence of real numbers has a convergent subsequence, it follows that $g(z) = 0$ in $\mathbb{C}$, by the identity theorem. So, $f(z)$ is real on the real axis. As $f(x)$ is real for $x \in \mathbb{R}$, we can apply the mean value theorem of calculus on the interval $[a_{2n+1}, a_{2n}]$. Thus, for each n, there exists a c_n such that

$$a_{2n+1} \leq c_n \leq a_{2n} \quad \text{and} \quad f'(c_n) = 0.$$

As $a_n \to 0$, we see that $c_n \to 0$. Consequently, $f'(z) = 0$ in $\mathbb{C}$ by the identity theorem.

Questions 8.72:

2. $f(z) = z^n$ on $|z| = r$.
3. Not necessarily. If $f(z) = e^z$, then $|e^{re^{i\theta_1}}| < |e^z|$ for some $|z| < r$ whenever $\theta_1 \neq 2k\pi$.
4. e^z on $\{z : \mathrm{Re}\, z < 0\} \cup \{0\}$ attains a maximum at $z = 0$.

Exercises 8.73:

3. For each $n \in \mathbb{N}$, let $f_n(z) = z^n f(z)$, and $F_n(z) = f_n(z)\overline{f_n(\overline{z})}$. Then, F_n is analytic and for $\epsilon > 0$ there exists an n such that $|F_n(\zeta)| < \epsilon$ for all $\zeta \in C$. By the Maximum modulus principle, this inequality yields that $|x^{2n} f(x)| \leq \epsilon$ for $x \in (0,1)$.
4. (a) Max at $z = r$, min $z = -r$.
   (b) Max and min everywhere.
   (c) Max at $z = r$ and min $z = ir$.
   (d) Max at $z = -r$, min $z = r$.
6. Observe that $|e^{z^2 - iz}| = e^{\mathrm{Re}\,(z^2 - iz)}$, and then maximize the quantity $\mathrm{Re}\,\{z^2 - iz\}$.
9. Use Theorem 8.38 and the fact that $|e^{P(z)}|$ is continuous.
12. Use Schwarz's lemma.
13. Suppose that such an f exists. Define

$$F(z) = \phi_{3/4} \circ f \circ \phi_{1/2}(z), \quad \phi_a(z) = \frac{a - z}{1 - \overline{a}z}.$$

Then F satisfies the hypothesis of the Schwarz lemma. But a computation gives that $F'(0) = 32/21$, which is a contradiction. Thus, we see that no such f can exist. See also Example 8.70. Define $F(z) = f(z/3)$. Then, F is analytic for $|z| \leq 1$ and $|F(z)| \leq 1$ for $|z| \leq 1$ so that $F(z)$ has zeros at $a_k = w_k/3$, $k = 0, 1, \dots, n-1$. Therefore, $F(z)$ is Blaschke product with zeros at a_k and so,

$$|F(0)| = \prod_{k=0}^{n-1} \left|\frac{w_k}{3}\right| = \frac{1}{3^n}.$$

Questions 9.8:

2. No. $\sum_{n=1}^{\infty}(z^n/n^2)$ converges uniformly on the annulus $\frac{1}{2} \leq |z| \leq 1$, but is not analytic on $|z| = 1$.
4. $\sum_{n=1}^{\infty}(1/n^z)$ is analytic in a half-plane.
5. Only for constant functions.
7. If $f(z)$ is analytic in an annulus, then the identity is valid in that annulus.

Exercises 9.9:

1. Use partial fractions.
4. $\displaystyle\sum_{n=0}^{\infty} \frac{z^{2n}}{n!} + \sum_{n=0}^{\infty} \frac{1}{z^{2n} n!}$.
5. (iii) $\displaystyle\frac{1}{a-b}\sum_{n=0}^{\infty} \frac{a^n - b^n}{z^{n+1}}$ (iv) $\displaystyle\sum_{n=0}^{\infty}(-1)^n \frac{(z-a)^{n-1}}{(a-b)^{n+1}}$.
6. We note that
$$\begin{aligned} f(z) &= \frac{1}{z}\left[\frac{1}{z-b} - \frac{1}{z-a}\right]\frac{1}{b-a} \\ &= \left[\left(\frac{1}{z-b} - \frac{1}{z}\right)\frac{1}{b} - \left(\frac{1}{z-a} - \frac{1}{z}\right)\frac{1}{a}\right]\frac{1}{b-a} \\ &= \frac{1}{ab(b-a)}\left[\frac{b-a}{z} - \frac{b}{z-a} + \frac{a}{z-b}\right] \end{aligned}$$
and the rest of the calculation is routine as in Example 9.7.
7. The calculation is routine once we write $f(z)$ as
$$f(z) = 1 + \frac{1}{z+1} + \frac{3}{4+z}.$$
8. (a) $\displaystyle -\frac{i}{4(z-i)}$ (d) $\displaystyle\frac{1}{z^3} - \frac{1}{6z}$.
9. (a) $\displaystyle\sum_{k=0}^{\infty} \frac{z^n}{k! z^k}$ (b) $\displaystyle\sum_{n=0}^{\infty} \frac{a_n}{z^n}$, where $\displaystyle a_n = \sum_{k=1}^{n} \binom{n-1}{k-1}\frac{1}{k!}$.
11. $\displaystyle\frac{4}{z^4} + \frac{4}{z^3} + \frac{8}{3z^2} + \frac{4}{3z}$.

Questions 9.21:

1. $1/\sin(1/z)$ has infinitely many singularities in the compact set $|z| \leq 1$. All but one, $z = 0$, are isolated.
2. A counterexample will be constructed in Chapter 12.
5. $e^z - (7 - 2i)$.
7. Only constant functions.
9. No, by definition.

Exercises 9.22:

4. (b) Simple poles at $z = 1/(2k+1)\pi i$, nonisolated essential singularity at $z = 0$.
   (e) Isolated essential singularities at $z = 0$ and $z = \infty$.
   (g) Branch point at $z = 1$ and simple pole at $z = -1$.
6. (a) Removable (b) Simple pole.
   (c) (d), (f) are all isolated essential singularities.
   (e) is a nonisolated essential singularity.
7. $f(z) = \dfrac{z - z_0}{(z - z_0)(z - z_1)^k} e^{1/(z-z_2)}$.
13. Set $f(z) = A/(1 - z/z_0) + F(z)$, where A is a constant and $F(z)$ is analytic for $|z| < R$.
14. Set $(1 + z)^{1/z} = e^{(1/z)\log(1+z)} = e^{(1/z)[\operatorname{Log}(1+z)+2k\pi i]}$.

Questions 9.36:

3. Morera's theorem cannot be applied because $\sin(1/z^2)$ is not continuous at $z = 0$.
5. None. The residue theorem is just a convenient form of Cauchy's theorem.
8. Because $1 + x^{2n+1}$ has a singularity on the real axis.

Exercises 9.37:

1. (a) At $z = \pi/2 + k\pi$, the residue is $(-1)^{k+1}$.
   (b) At $z = 1$, the residue is -2; at $z = 2$, the residue is 2.
   (c) At $z = 0$, the residue is $\begin{cases} 0 & \text{if } n \text{ is even,} \\ -\dfrac{1}{(n+1)!} & \text{if } n = 4k+1, \\ \dfrac{1}{(n+1)!} & \text{if } n = 4k+3. \end{cases}$
3. As $f'(a) \neq 0$, $f(z) - f(a) \neq 0$ in a deleted neighborhood of a. Therefore,
   $$\lim_{z \to a} (z - a) \frac{1}{f(z) - f(a)} = \frac{1}{f'(a)} \neq 0$$
   and the conclusion is a consequence of the Residue theorem.
5. (a) 0 (b) $2\pi i(1 + 2e + 2e^4)$.

6. We may rewrite the given integral as

$$I = \frac{1}{2\pi i}\int_{|z|=2} f(z)\,dz, \quad f(z) = \frac{z^n}{(z-e^{i\phi})(z-e^{-i\phi})},$$

The poles of $f(z)$ are at $z = e^{i\phi}$, $e^{-i\phi}$ and both lie inside $|z| = 2$. It follows easily that

$$I = \operatorname{Res}\,[f(z); e^{i\phi}] + \operatorname{Res}\,[f(z); e^{-i\phi}] = \frac{\sin n\phi}{\sin\phi}.$$

7. (a) $(2-4n)i$ (b) $(2+4n)i$.
9. Note that $f(z) = 0$ implies that $z \in \{*\sqrt{-3+2i}, *\sqrt{-3-2i}\}$, where

$$*\sqrt{-3+2i} = \{\pm a\} \quad\text{and}\quad *\sqrt{-3-2i} = \{\pm\overline{a}\},$$

where

$$a = \sqrt{\frac{-3+\sqrt{13}}{2}} + i\sqrt{\frac{3+\sqrt{13}}{2}}.$$

The poles lying in the upper half-plane are a and $b := -\overline{a}$. They are simple poles. Therefore we find that

$$\operatorname{Res}\left[\frac{z^2}{f(z)}; a\right] = \frac{a^2}{f'(a)} = \frac{a}{4(a^2+3)} = -\frac{ia}{8}$$

as $a^2 = -3+2i$. Similarly we have

$$\operatorname{Res}\left[\frac{z^2}{f(z)}; b\right] = \frac{b}{4(b^2+3)} = \frac{b}{4[(-3-2i)+3]} = \frac{ib}{8}.$$

10. Consider the polynomial equation $f(z) = a_0 + a_1 z + \cdots + z^n$. Since $|f(z)| \to \infty$ as $z \to \infty$, for sufficiently large R, we have $|f(z)| > 0$ for $|z| \geq R$. If we let $F(z) = f'(z)/f(z)$ and $C = \{z : |z| = R\}$, described in the positive direction, then

$$\begin{aligned}
\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz = \frac{1}{2\pi i}\int_C F(z)\,dz &= -\operatorname{Res}\,[F(z);\infty] \\
&= \operatorname{Res}\left[\frac{F(1/z)}{z^2}; 0\right] \\
&= \lim_{z\to 0} z\left[\frac{f'(1/z)}{z^2 f(1/z)}\right] \\
&= n.
\end{aligned}$$

12. Keep c fixed and let $R \to \infty$.

Questions 9.58:

1. $e^{1/z}$, $z \neq 0$.
6. $|z|$ maps the plane onto the ray $\mathrm{Re}\, u \geq 0$, $v = 0$, which is not open.
7. e^z maps the plane onto the punctured plane, which is not closed.
8. Consider the exponential function $f(z) = e^z$.

Exercises 9.59:

3. The sum of the roots of $P(z)$.
4. If $f(z) = a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n$, $n \geq 1$ and R is chosen large enough so that $|f(z)| \geq 1$ for all $|z| \geq R$ then, for $|z| \geq R$, we have

$$\frac{f'(z)}{f(z)} = \frac{nz^{n-1} + \cdots}{z^n + \cdots} = \frac{n}{z} + \text{ terms in } \frac{1}{z^k},\ k \geq 2.$$

Thus,

$$\frac{1}{2\pi i} \int_{|z|=R} \frac{f'(z)}{f(z)}\, dz = \frac{n}{2\pi i} \int_{|z|=R} \frac{dz}{z} = n.$$

Since f has no poles in $\mathbb{C}$, the fundamental theorem of algebra follows.

10. For $|z| = 1$, $z = x + iy$, $|-az^n| = a > e > e^x = |e^z|$.
12. If $p(z) = z^3 + iz + 1$, then

$$p(x) = 0 \implies x^3 + 1 = 0 \text{ and } x = 0$$

which is not possible. Similarly,

$$p(iy) = 0 \implies 1 - y = 0 \text{ and } y^3 = 0$$

which is again not possible.

13. For $|z| = 1$,

$$|z^4 + 1| \leq |z|^4 + 1 = 2 < |-6z| = 6$$

and for $|z| = 2$,

$$|-6z + 1| \leq 6|z| + 1 = 13 < |z|^4 = 2^4.$$

16. For $|z| = 3/2$,

$$|z^3 + 1| \leq |z|^3 + 1 = 27/8 + 1 < |z|^4 = 81/16$$

and for $|z| = 3/4$,

$$|z|^4 = 81/256 \leq -|z|^3 + 1 = -27/64 + 1 \leq |z^3 + 1|.$$

18. Use Corollary 9.47.
19. Use Hurwitz's theorem.

Chapter 10:

Questions 10.15:

5. Not according to Picard's theorem.
8. $u(z) = \begin{cases} x+y \text{ if } |z|<1, \\ 1 \quad\quad \text{if } |z|=1. \end{cases}$
11. Theorem 10.14 generalizes Theorem 8.35 because $\operatorname{Re} f(z) \le |f(z)|$, and reduces to Theorem 10.4 when $\lambda = 0$.
14. By the mean value theorem, $u(2,-1) = -2$.
15. $u(z) = ay + b$ for some constants a and b.
18. As $f = u + iv$ is analytic in D, $\operatorname{Im}(f^2) = 2uv$ is harmonic. Also u^2 is harmonic on D iff u is constant.
19. No.

Exercises 10.16:

10. $u = 1$.
11. $u = y$.
13. Use Cauchy's integral formula in conjunction with Theorem 10.13.

Questions 10.34:

2. Yes, by finding an upper bound on the entire function using (10.14), and then applying Theorem 8.35.
3. The proof of uniqueness is easy in the case when F is continuous on $|z| = R$. Then u would be harmonic for $|z| < R$, continuous on $|z| = R$, and equal to F on $|z| = R$. Let u_1 be another such function. Then $u - u_1 = 0$ on $|z| = R$ showing that $u = u_1$ for $|z| < R$ (by Corollary 10.10).
4. In Chapter 11, we shall discuss mappings from the disk to other domains. This will enable us to solve the Dirichlet problem for other domains.
5. Theorem 10.6 gives the value of the harmonic function at the center of the circle, whereas Theorem 10.18 gives the value for all points inside.

Exercises 10.35:

6. $u(re^{i\theta}) = \dfrac{1}{\pi}\tan^{-1}\dfrac{1-r^2}{2r\sin\theta} \quad (0 \le \tan^{-1} t \le \pi)$.
8. Show that
$$\tan^{-1}\frac{r\sin\theta}{1+r\cos\theta} = \operatorname{Im}\log(1+z) \quad (z = re^{i\theta}).$$
9. With $C = [-R, R] \cup \Gamma_R$, where Γ_R is the upper semi-circular contour from R to $-R$, we write
$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\int_C f(\zeta)\left(\frac{1}{\zeta - z} - \frac{1}{\zeta - \overline{z}}\right) d\zeta \\ &= \frac{y}{\pi}\int_{-R}^{R} \frac{f(t)\,dt}{(t-z)(t-\overline{z})} + \frac{y}{\pi}\int_{\Gamma_R} \frac{f(\zeta)\,d\zeta}{(\zeta - z)(\zeta - \overline{z})}. \end{aligned}$$

As usual the second integral can be shown to approach zero as $R \to \infty$. Allowing $R \to \infty$ and noting $f(t) = u(t,0) + iv(t,0)$, the result follows by equating real and imaginary parts.

11. Use (10.22) by noting that

$$\frac{\zeta + z}{(\zeta - z)\zeta} = \frac{-1}{\zeta} + \frac{2}{\zeta - z}.$$

12. See Example 10.31.
13. Following the idea of Example 10.31, consider

$$u(x,y) = \mathrm{Im}\,[a\,\mathrm{Log}\,(z-1) + b\,\mathrm{Log}\,(z+1) + c]$$

which is harmonic for $\mathrm{Im}\,z > 0$. Use the given conditions to find the constants a, b, c.
Note: More generally, one can solve the following Dirichlet problem: $u_{xx} + u_{yy} = 0$ for $\mathrm{Im}\,z > 0$ subject to the boundary conditions

$$u(x,0) = a_k \quad \text{for} \quad x_k < x < x_{k+1}, \quad k = 0, 1, \ldots, n,$$

where $x_0 = -\infty$ and $x_{n+1} = \infty$.

14. According to the Poisson integral formula, for $\zeta = e^{i\phi}, z = re^{i\theta}$, the function

$$u(z) = \frac{1-r^2}{2\pi}\int_0^{\pi} \frac{1}{1 - 2r\cos(\theta - \phi) + r^2}\, d\phi = \frac{1}{2\pi}\int_0^{\pi} \mathrm{Re}\,\frac{\zeta + z}{\zeta - z}\, d\phi$$

is harmonic for $|z| < 1$. An integration shows that

$$\begin{aligned} u(z) &= \frac{1}{\pi}\tan^{-1}\left(\frac{1+r}{1-r}\tan\frac{\phi - \theta}{2}\right)\Bigg|_0^{\pi} \\ &= \frac{1}{\pi}\left[\tan^{-1}\left(\frac{1+r}{1-r}\tan\frac{\pi - \theta}{2}\right) - \tan^{-1}\left(\frac{1+r}{1-r}\tan\frac{-\theta}{2}\right)\right]. \end{aligned}$$

From the trigonometric identities

$$\begin{aligned} \tan(\alpha - \beta) &= \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha\tan\beta}, \\ \tan\left(\frac{\pi - \theta}{2}\right) &= \cot\frac{\theta}{2}, \\ \tan\frac{\theta}{2} + \cot\frac{\theta}{2} &= \frac{2}{\sin\theta}, \end{aligned}$$

it follows easily that

$$\tan \pi u(z) = -\frac{1 - r^2}{2r\sin\theta}.$$

Choosing the determination of the inverse tangent whose values lie in the interval $[0, \pi]$, we have

$$u(re^{i\theta}) = \frac{1}{\pi} \tan^{-1}\left(-\frac{1-r^2}{2r\sin\theta}\right).$$

With this determination,

$$\lim_{r\to 1} \tan^{-1}\left(-\frac{1-r^2}{2r\sin\theta}\right) = \begin{cases} \pi & \text{if } 0 < \theta < \pi \\ 0 & \text{if } \pi < \theta < 2\pi \end{cases}$$

the desired boundary conditions are satisfied.

15. Use for example, (10.14).

Questions 10.45:

2. In view of Theorem 10.38, the mean-value property does not hold, whenever the product of two harmonic functions is not harmonic.
4. Yes, just consider $\{-u_n(z)\}$.
5. Consider $|z| < R$ and $\operatorname{Re} f(z) > \alpha$.
7. We used the fact that $\int_0^{2\pi} |u(re^{i\theta})|\, d\theta = \int_0^{2\pi} u(re^{i\theta})\, d\theta$.

Exercises 10.46:

5. Set $g(z) = (1-\alpha)f(z) + \alpha$, where $\operatorname{Re} f(z) > 0$. Then apply Theorem 10.42. Why can't α exceed 1?
7. Set $f(z) = (g(z) - \alpha)/(1-\alpha)$, and apply Theorem 10.44.

Questions 11.7:

1. By convention the line itself is the tangent line.
2. Only at the point of intersection.
3. Not if the partial derivatives are continuous. See Nehari [N].
9. A one-to-one map is conformal if it is analytic; a conformal map is locally one-to-one.
12. Only the composition.

Exercises 11.8:

1. $f(z) = e^{3\pi i(z-z_0)/\epsilon}$.

Questions 11.16:

2. No. Even the family of constant polynomials is not.
3. $\{z^n\}$ is uniformly bounded on $|z| < 1$, but $\{nz^{n-1}\}$ is not.
4. No. Let F consist of one function $f(z)$ defined by

$$f(z) = \begin{cases} 1 \text{ if } z \neq 1/n, \\ n \text{ if } z = 1/n. \end{cases}$$

This is unbounded in every neighborhood of the origin.

5. We have assumed that the line segment between any two points lies in the domain. Thus the proof is valid for any convex domain.
6. A sequence must be countable.

Questions 11.30:

2. $f(z) = az$, $a > 0$, maps the annulus $r < |z| < R$ conformally onto the annulus $ar < |w| < aR$.
5. No, because the punctured disk is not simply connected.
13. In the construction of the analytic square-root function.
14. In order to ensure that the normal family constructed is nonempty.
19. A desired map is given by $\phi(z) = e^{\pi z/\alpha}$.
20. A desired map is given by $f(z) = (e^{iz} - 1)/(e^{iz} + 1)$.

Exercises 11.31:

2. Suppose the plane were conformally equivalent to a simply connected domain D other than itself. Since D is conformally equivalent to a bounded domain, there would have to be an entire function mapping onto a bounded domain.
6. $f(z) = \dfrac{r_2}{r_1} z$.

Questions 11.43:

2. Theorem 11.32 enables us to prove theorems about $\mathcal{S}$ from theorems about $\mathcal{T}$.
4. Functions of the form $z/(1 - e^{i\alpha}z)^2$ are unbounded.
6. This follows from the fact that if $J(f)$ is a continuous functional defined on a compact family $\mathcal{F}$, then the problem $|J(F)| = \max$ has a solution for some $f \in \mathcal{F}$.

Exercises 11.44:

1. $1/z$.
4. Its derivative is 0 at $z = -\frac{1}{2}e^{-i\alpha}$.
7. Consider $f(z_1) - f(z_0) = (z_1 - z_0) + \sum_{n=2}^{\infty} a_n(z_1^n - z_0^n)$.
8. If $\sum_{n=2}^{\infty} n|a_n| > 1$, show that $f'(r_0) = 0$ for some $0 < r_0 < 1$.

Questions 12.16:

3. See Example 12.11 and Exercise 12.17 (6).
4. It can approach ∞, 0, or oscillate.
10. The sequence $\{a_n\} \to 0$. The series $\sum_{n=0}^{\infty} a_n$ does not converge absolutely. It may or may not converge.
13. It is an entire function.

Exercises 12.17:

3. Set $-\operatorname{Log}(1 - a_n) = a_n + a_n^2(1/2 + a_n/3 + \cdots)$, and apply Theorem 12.5.

5. Apply Exercise 3.
8. (a) $|z| < 1$ (b) $|z| < 1$ (c) $\mathrm{Re}\, z > 1$.

Questions 12.29:

1. This will be established in the next section.
7. It is necessary that the convergence be absolute. For instance,

$$\prod_{n=1}^{\infty}\left(1+\frac{(-1)^{n+1}z}{\sqrt{n}}\right) \qquad \text{diverges at } z=1.$$

8. The series expansion can be determined from the product expansion, but the converse is not true.

Exercises 12.30:

1. For example, $\prod_{n=1}^{\infty}\left(1-\frac{1}{n(R+z)}\right)e^{1/(R+z)n}$.
2. $\displaystyle\prod_{n=2}^{\infty}\left(1-\frac{z}{\ln n}\right)e^{(z/\ln n)+(1/2)(z/\ln n)^2+\cdots+(1/n)(z/\ln n)^n}+\cdots.$
3. (b) $f(z)=\displaystyle\prod_{n=1}^{\infty}(1-z/n)^n e^{z^2/2n}$.
4. (a) Use the product expansion for $\sin\pi z$. and note that the value is $(e^{\pi}-e^{-\pi})/(2\pi)$.
5. (b) $g(z) = z\,\mathrm{Log}\,(-2i)$.
11. Use the identity $\cos z = (\sin 2z)/(2\sin z)$.
13. Using the series expansion of $e^{-z/n}$, it follows that

$$\left(1+\frac{z}{a+n}\right)e^{-z/n}=\sum_{k=0}^{\infty}(-1)^k\frac{n(1-k)+a}{(n+a)k!n^k}z^k := 1+a_n(z).$$

We observe that $\sum_{k=0}^{\infty}|a_n(z)|$ converges for all z, because

$$\lim_{n\to\infty}\left|\frac{a_n(z)}{1/n^2}\right| = |az|.$$

Thus, the given product represents an entire function.

14. $\displaystyle\frac{9}{8}\,\frac{e^{\pi^2}+e^{-\pi^2}-(e^{\pi^2/3}+e^{-\pi^2/3})}{\pi^4}.$
15. (b) First set $z=\frac{1}{2}$ and then $z=\frac{1}{4}$. Now divide the latter expression by the former.

    (c) Same as above, with $z=\frac{1}{3}$ and $z=\frac{1}{6}$.
16. Logarithmic differentiation of $\Gamma(z)$ defined by (12.19) gives

$$\frac{\Gamma'(z)}{\Gamma(z)}=-\frac{1}{z}-\gamma+z\sum_{n=1}^{\infty}\frac{1}{n(z+n)}=-\frac{1}{z}-\gamma-\sum_{n=1}^{\infty}\left(\frac{1}{z+n}-\frac{1}{n}\right),$$

for $z\in\mathbb{C}\backslash\{0,-1,-2,\ldots\}$. Another differentiation yields the formula.

18. Use the identity $e^z - 1 = 2ie^{z/2}\sin(z/2i)$.

Questions 12.39:

2. They are meromorphic.
5. In order to obtain a Maclaurin expansion.
6. No, since the entire function $g(z)$ constructed in the proof of Theorem 13.8 is not unique.

Exercises 12.40:

1. $f(z) = \displaystyle\sum_{n=1}^{\infty} \frac{1}{(z-n)^n}$.
2. $\dfrac{i}{2}e^{-iz}$.
8. Differentiation of the partial fraction decomposition of $\pi\cot(\pi z)$ gives the desired result.
10. Take logarithmic derivatives of both sides in the identity proved in Exercise 12.30 (18).
11. Define $P_n(z) = \prod_{k=1}^{n}\left(1+h^{2k-1}e^z\right)\left(1+h^{2k-1}e^{-z}\right)$ and

$$Q_n(z) = \prod_{k=1}^{n}\left(1+h^{2k-1}e^{z+2\log h}\right)\left(1+h^{2k-1}e^{-(z+2\log h)}\right).$$

Then

$$\begin{aligned}\frac{Q_n(z)}{P_n(z)} &= \frac{\prod_{k=1}^{n}\left(1+h^{2k+1}e^z\right)\left(1+h^{2k-3}e^{-z}\right)}{\prod_{k=1}^{n}\left(1+h^{2k-1}e^z\right)\left(1+h^{2k-1}e^{-z}\right)}\\ &= \frac{1+h^{2n+1}e^z}{1+he^z}\,\frac{1+h^{-1}e^{-z}}{1+h^{2n-1}e^{-z}} \;\to\; \frac{1+h^{-1}e^{-z}}{1+he^z} = \frac{1}{he^z}\end{aligned}$$

as $n\to\infty$ which confirms the truth of the functional equation.

Questions 13.11:

1. Only if $\mathcal{D}_0 \cap \mathcal{D}_1 = \emptyset$.
2. $\sin\dfrac{1}{1-z_n} = 1$ when $z_n = 1 - \dfrac{1}{n\pi}$.
8. When their intersection is nonempty.
10. It can be shown that if

$$f(z) = \sum_{n=1}^{\infty} a_{n_k} z^{n_k} \quad \text{with} \quad n_{k+1} > (1+\epsilon)n_k \quad (\epsilon > 0),$$

then the circle of convergence of the series is a natural boundary of the function.

Exercises 13.12:

1. $f(z) = \dfrac{1}{\prod_{k=1}^{n}(z-e^{i\theta_k})}$.

2. Write

$$\frac{1}{1-z} = \frac{1}{1+p-(z+p)} = \sum_{n=0}^{\infty} \frac{(z+p)^n}{(1+p)^{n+1}}, \quad |z+p| < 1+p,$$

and set, for example,

$$f_{p+1}(z) = \sum_{n=0}^{\infty} \frac{(z+p)^n}{(1+p)^{n+1}} \text{ with } D_{p+1} = \{z : |z+p| < 1\}$$

where $p = 0, 1, 2, \ldots$.

7. Expand $g(z) = f(z)/(1-z)$ in a series, and show that $(1-z)g(z) \to \infty$ as $z \to 1$ through real values.
8. Set $f(z) = \dfrac{\alpha_1}{z-e^{i\theta_1}} + \dfrac{\alpha_2}{z-e^{i\theta_2}} + \cdots + \dfrac{\alpha_k}{z-e^{i\theta_k}} + \sum_{n=0}^{\infty} b_n z^n$, where $\sum_{n=0}^{\infty} b_n z^n$ is analytic for $|z| \leq 1$.
10. Apply Theorem 13.8 for $f(-z)$.

Questions 13.20:

1. As $e^{-t} > e^{-1}$ for all $0 < \delta \leq t \leq 1$,

$$\int_\delta^1 t^{x-1} e^{-t}\, dt > \frac{1}{e} \int_\delta^1 t^{x-1}\, dt = \frac{1}{e}\left(\frac{1-\delta^x}{x}\right)$$

which approaches ∞ as $x \to 0^+$, for each $\delta > 0$.

8. We need the uniform convergence of the sequence $\{(1-t/n)^n\}$. This sequence does not converge uniformly on the line.
10. The series $\sum_{n=1}^{\infty} a_n z^n$ is analytic in a disk and converges uniformly on compact subsets of the disk; it can be shown that $\sum_{n=1}^{\infty}(a_n/n^z)$ is analytic in a half-plane and converges uniformly on compact subsets of the half-plane.
11. Entire.

Exercises 13.21:

1. First separate the product for $\Gamma(2z)$ into even and odd terms:

$$\frac{1}{\Gamma(2z)} = 2ze^{2\gamma z} \prod_{k=1}^{\infty}\left(1+\frac{z}{k}\right)e^{-z/k} \prod_{k=0}^{\infty}\left(1+\frac{z}{k+1/2}\right)e^{-z/(k+1/2)}.$$

Deduce that $\Gamma(2z)/[\Gamma(z)\Gamma(z+1/2)]$ has the form ae^{bz}. Finally, evaluate a and b by setting $z = 1/2$ and $z = 1$. One can also use (13.7) to prove this formula.

3. On the line $\operatorname{Re} s = 2$, show that $\operatorname{Re}\zeta(s) > 1 - \sum_{n=2}^{\infty} 1/n^2$.
6. Use the identities $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$ and $\sin 2\theta = 2\sin\theta\cos\theta$ in (13.32).

7. By (13.18), we have

$$\frac{1}{n^{1/4}} = \frac{1}{\Gamma(1/4)} \int_0^\infty t^{(1/4)-1} e^{-nt}\, dt$$

so that

$$\sum_{n=1}^\infty \frac{z^n}{n^{1/4}} = \frac{1}{\Gamma(1/4)} \int_0^\infty t^{-3/4} \sum_{n=1}^\infty (e^{-t}z)^n\, dt = \frac{1}{\Gamma(1/4)} \int_0^\infty \frac{z}{t^{3/4}(e^t - z)}\, dt.$$

The integral on the right defines an analytic function outside of the interval $[1, \infty)$.

10

Harmonic Functions

In Chapter 5, we saw that if an analytic function has a continuous second derivative, then the real (or imaginary) part of the function is harmonic. In Chapter 8, it was shown that *all* analytic functions are infinitely differentiable and in particular, have continuous second derivatives. Thus, the real part of an analytic function is always harmonic.

In this chapter, we examine the extent to which the converse is true. In simply connected domains, we show that every harmonic function is the real part of *some* analytic function. This result enables us to prove several theorems for harmonic functions that are analogous to theorems for analytic functions. In particular, an analog to Cauchy's integral formula, known as Poisson's integral formula, gives a method for determining the values of a harmonic function inside a disk from the behavior at its boundary points.

10.1 Comparison with Analytic Functions

Recall that a continuous real-valued function $u(x, y)$, defined and single-valued in a domain D, is said to be harmonic in D if it has continuous first and second partial derivatives that satisfy Laplace's equation

$$u_{xx} + u_{yy} = 0.$$

In Section 5.3, we illustrated how the Cauchy–Riemann equations might be used to construct a function $v(x, y)$ conjugate to a given harmonic function $u(x, y)$; that is, a function $v(x, y)$ was found for which $f(z) = u(x, y) + iv(x, y) = u(z) + iv(z)$ was analytic. The method entailed finding all functions $v(z)$ satisfying the two conditions

$$u_x = v_y, \quad u_y = -v_x.$$

This method was successful when the partial integration $\int v_y \, dy$ could explicitly be solved. We now give general conditions for the existence of an analytic function whose real part is a prescribed harmonic function. First note that in view of the Cauchy–Riemann equations, the derivative of any analytic function $f(z) = u(z) + iv(z)$ may be expressed as

$$f'(z) = u_x(z) - iu_y(z).$$

Hence we can find f (by integration) directly from u. The details follow.

Theorem 10.1. *If u is harmonic on a simply connected domain D, then there exists an analytic function on D whose real part equals u.*

Proof. Set $g(z) = u_x(z) - iu_y(z) := U(z) + iV(z)$, $z \in D$. Then by Laplace's equation,

$$U_x - V_y = u_{xx} - (-u_{yy}) = 0. \tag{10.1}$$

Since the mixed partial derivatives of $u(z)$ are continuous in D,

$$U_y + V_x = (u_x)_y + (-u_y)_x = 0. \tag{10.2}$$

But (10.1) and (10.2) are the Cauchy–Riemann equations for $g = U + iV$. Noting that U_x, U_y, V_x, V_y are all continuous, we may apply Theorem 5.17 to establish the analyticity of $g(z)$ in D.

Next choose any point z_0 in D, and set

$$F(z) = \int_{z_0}^{z} g(\zeta)\, d\zeta.$$

Then, by Corollary 8.15, $F(z)$ is analytic in D with

$$F'(z) = g(z) = u_x(z) - iu_y(z).$$

Observe that the derivative of $F(z)$ may also be expressed as

$$F'(z) = \frac{\partial}{\partial x}\operatorname{Re} F(z) - i\frac{\partial}{\partial y}\operatorname{Re} F(z).$$

Thus $u(z)$ and $\operatorname{Re} F(z)$ have the same first partial derivatives in D, so that

$$\operatorname{Re} F(z) = u(z) + c \quad (c \text{ a real constant}).$$

Hence, the function

$$f(z) = F(z) - c = \int_{z_0}^{z} (u_x(\zeta) - iu_y(\zeta))\, d\zeta - c$$

is analytic in D with $\operatorname{Re} f(z) = u(z)$. ∎

Corollary 10.2. *If u is harmonic on a simply connected domain D, then there exists an analytic function on D whose imaginary part equals u.*

Proof. By Theorem 10.1, there exists an analytic function $h(z)$ such that $\operatorname{Re} h(z) = u(z)$. But then $f(z) = ih(z)$ is analytic with $\operatorname{Im} f(z) = \operatorname{Re} h(z) = u(z)$. ∎

Example 10.3. Let $u(x, y) = \sin x \cosh y + \cos x \sinh y + x^2 - y^2 + 2xy$. It can be easily seen that u is harmonic in $\mathbb{C}$. Following the proof of Theorem 10.1,

$$\begin{aligned} f'(z) = u_x - iu_y = {} & \cos x \cosh y - \sin x \sinh y + 2x + 2y \\ & -i\,(\sin x \sinh y + \cos x \cosh y - 2y + 2x). \end{aligned}$$

As $\cos(iy) = \cosh y$ and $-i \sin(iy) = \sinh y$, we can simplify the last equation and obtain

$$f'(z) = (1 - i)(\cos z + 2z).$$

Thus, $f(z) = (1 - i)(\sin z + z^2) + c$. ●

The requirement in Theorem 10.1 that the domain be simply connected is essential. For example, the function

$$u(z) = u(x, y) = \ln \sqrt{x^2 + y^2} = \ln |z|$$

is harmonic in the punctured plane $\mathbb{C} \backslash \{0\}$. Each point in $\mathbb{C} \backslash \{0\}$ has a neighborhood where $\log z$ has a single-valued analytic branch. In other words, we say that $u(z)$ is locally the real part of an analytic function as guaranteed by Theorem 10.1. Therefore, $u(z) = \ln |z|$, being the real part of an analytic function, is harmonic in such neighborhoods. We also know that the principal logarithm $\operatorname{Log} z$ defined by

$$\operatorname{Log} z = \ln |z| + i \operatorname{Arg} z$$

is analytic in the cut plane $D = \mathbb{C} \backslash (-\infty, 0]$. Now if some function

$$f(z) = \ln |z| + iv(z)$$

were analytic throughout the punctured plane $\mathbb{C} \backslash \{0\}$, then g defined by

$$g(z) = f(z) - \operatorname{Log} z$$

would be analytic in the slit plane $D = \mathbb{C} \backslash (-\infty, 0]$. Since $g(z)$ is purely imaginary in D, an application of the Cauchy–Riemann equations shows that $g(z)$ must be constant in D. Thus, any function analytic in D whose real part is $\ln |z|$ must be of the form

$$u(z) + iv(z) = \operatorname{Log} z + ic,$$

where c is a real constant. It follows that $v(z) = \operatorname{Arg} z + c$. But then

$$\lim_{\substack{y \to 0 \\ y > 0}} v(-1 + iy) = \lim_{\substack{y \to 0 \\ y > 0}} \operatorname{Arg}(-1 + iy) + c = \pi + c$$

and

$$\lim_{\substack{y\to 0\\ y<0}} v(-1+iy) = \lim_{\substack{y\to 0\\ y<0}} \operatorname{Arg}(-1+iy) + c = -\pi + c$$

which means that v is discontinuous at -1, a contradiction. An argument similar to this shows that v is not continuous at all points in the negative real axis $(-\infty, 0]$. Thus, there is no hope for defining an analytic function in $\mathbb{C}\setminus\{0\}$ whose real part is $u(z) = \ln|z|$. Hence, *a harmonic function need not have an analytic completion in a multiply connected domain.*

In view of Theorem 10.1, we may now modify some of the theorems in Chapter 8 to obtain harmonic analogs. Our next theorem is the harmonic analog of Liouville's theorem.

Theorem 10.4. *A function harmonic and bounded in $\mathbb{C}$ must be a constant.*

Proof. Suppose $u(z)$ is harmonic and bounded in the plane. Theorem 10.1 guarantees the existence of an entire function $f(z)$ whose real part is $u(z)$. But then

$$g(z) = e^{f(z)}$$

is an entire function too. Since $|g(z)| = e^{u(z)}$, $g(z)$ is also bounded in the plane. By Liouville's theorem $g(z)$, and hence $u(z) = \ln|g(z)|$, must be constant. ■

Clearly, Theorem 10.4 may be restated in a general form as follows:

Theorem 10.5. *If the real or imaginary part of an entire function is bounded above or below by a real number M, then the function is a constant.*

We now prove an analog to Gauss's mean-value theorem for analytic functions. This is one of the fundamental facts about harmonic functions, called the mean value property of harmonic functions.

Theorem 10.6. (Mean Value Property) *Suppose $u(z)$ is harmonic in a domain containing the disk $|z - z_0| \le R$. Then*

$$u(z_0) = \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + Re^{i\theta})\, d\theta.$$

Proof. Let $f(z)$ be a function analytic in $|z - z_0| \le R$ whose real part is $u(z)$. By Gauss's mean-value theorem,

$$f(z_0) = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + Re^{i\theta})\, d\theta.$$

The result follows upon taking real parts of both sides. ■

The right-hand side of the last formula gives in particular that the mean (or average) value u on the circle $|z - z_0| = R$ is simply the value of u at the center of the circle $|z - z_0| = R$. In Section 10.2, we shall consider a similar expression for a point of the disk $|z - z_0| < R$ other than the center. We have shown that the behavior of a harmonic function on the boundary of a closed and bounded region determines the behavior of the harmonic function throughout the region. For instance, a harmonic function u in the unit disk $|z| < 1$ that extends continuously to $|z| \leq 1$ is completely determined by its values on the boundary $|z| = 1$. The explicit formula for the value of u for each point in $|z| < 1$ is given by the Poisson integral formula for a harmonic function and this is the subject of the discussion in Section 10.2. Unlike the situation for analytic functions, this result cannot be improved to an arbitrary sequence of points in the region. For instance, the nonconstant function $u(z) = x$ is harmonic in the plane with $u(z) \equiv 0$ on the imaginary axis. Hence, "analytic" cannot be replaced with "harmonic" in the statement of Theorem 8.47. That is, even if $u(z)$ is harmonic in a domain D, $u(z_n) \equiv 0$, and $z_n \to z_0$ in D, we are not guaranteed that $u(z) \equiv 0$ in D. Thus, the analog of the identity principle (see Theorem 8.48) for analytic functions does not hold for harmonic functions. However, we can salvage the following:

Theorem 10.7. *If $u(z)$ is harmonic in a domain D and constant in the neighborhood of some point in D, then $u(z)$ is constant throughout D.*

Proof. Let A be the set of all points z_0 in D for which $u(z)$ is constant in some neighborhood of z_0. Clearly A is a nonempty open set. To prove that $A = D$, it suffices to show that $B = D \setminus A$ is open, for then B would have to be empty in order for D to be connected.

Suppose B is not open. Then for a point z_0 in B and an $\epsilon > 0$ there is a point z_1 in A such that $z_1 \in N(z_0; \epsilon) \subset D$. Since A is open, we can find a $\delta > 0$ sufficiently small so that $N(z_1; \delta) \subset N(z_0; \epsilon) \cap A$. Now construct an analytic function $f(z)$ such that

$$\operatorname{Re} f(z) = u(z) \quad \text{for all } z \text{ in } N(z_0; \epsilon).$$

Since $u(z)$ is constant in $N(z_1; \delta)$, $f'(z) = 0$ for z in $N(z_1; \delta)$. An application of Theorem 8.47 to $f'(z)$ shows that $f'(z) \equiv 0$ throughout $N(z_0; \epsilon)$. Then, by Theorem 5.9, $f(z)$ is constant in $N(z_o; \epsilon)$. Hence, $u(z) = \operatorname{Re} f(z)$ is also constant in $N(z_0; \epsilon)$, contradicting the assumption that $z_0 \in B$. ■

Example 10.8. Suppose that $u(z)$ is harmonic in a domain D such that the set $\{z \in D : u_x(z) = 0 = u_y(z)\}$ has a limit point in D. Then we can easily show that $u(z)$ is a constant throughout D.

To see this, we define

$$F(z) = u_x(z) - iu_y(z), \quad z \in D.$$

Then F is analytic in D and the set $\{z \in D : F(z) = 0\}$ has a limit point in D. By the uniqueness theorem for analytic functions (see Theorem 8.47), $F(z) \equiv 0$ in D and so, $u_x(z) = 0 = u_y(z)$ on D, i.e., u is a constant. ■

Analogous to the maximum and minimum modulus theorems for analytic functions are the maximum and minimum principles for harmonic functions. The fact that a harmonic function is locally the real part of an analytic function produces a number of important results. One of them is the maximum principle.

Theorem 10.9. (Maximum Principle for Harmonic Functions) *A nonconstant harmonic function cannot attain a maximum or a minimum in a domain.*

Note that a harmonic function $u(z)$ attains a maximum at a point z_0 if and only if the harmonic function $-u(z)$ attains a minimum at z_0. So the minimum principle can be derived directly from the maximum principle. This result has several proofs.

Proof. The maximum modulus theorem for analytic functions is a direct consequence of Gauss's mean-value theorem and the fact that an analytic function is continuous. Similarly, we may deduce the maximum principle for harmonic functions from the mean-value principle for harmonic functions (Theorem 10.6). Indeed, we assume that $u(z)$ attains the maximum at $z_0 \in D$. Then, for each r with $0 < r \leq \operatorname{dist}(z_0, D)$, Theorem 10.6 gives

$$\frac{1}{2\pi}\int_0^{2\pi} (u(z_0) - u(z_0 + Re^{i\theta}))\, d\theta = 0.$$

Since $u(z_0) - u(z_0 + Re^{i\theta})$ is a continuous function of θ and is nonnegative, we have

$$u(z_0) = u(z_0 + Re^{i\theta}) \quad \text{for } 0 \leq \theta \leq 2\pi.$$

Thus, $u(z) = u(z_0)$ for all z in some neighborhood $N(z_0; \delta)$. Hence, $u(z) = u(z_0)$ on D (see Theorem 10.7).

For a second proof, we assume that $u(z)$ is a nonconstant function harmonic in a domain D. Given z_0 in D, construct a function $f(z) = u(z) + iv(z)$ that is analytic in some neighborhood $N(z_0; \delta)$ of z_0.

We set $g(z) = e^{f(z)}$, and note that $|g(z)| = e^{u(z)}$. If z_0 were a maximum for $u(z)$ in this neighborhood, then z_0 would be a maximum for $|g(z)|$. By the maximum modulus theorem for analytic functions, the function g must be constant on $N(z_0; \delta)$. Therefore, u is constant on $N(z_0; \delta)$ and hence on D, which contradicts the assumption that u is nonconstant. The proof is complete.

Alternatively, one could use the open mapping theorem (Theorem 9.55). Then it follows that there exists an $\epsilon > 0$ such that $N(f(z_0); \epsilon)$ is contained in the image of $N(z_0; \delta)$ under $f(z)$. In particular, for some point $z_1 \in N(z_0; \delta)$ we have $\operatorname{Re} f(z_1) = u(z_0) + \epsilon/2$. Thus, z_0 is not a maximum of $u(z)$ in D. ■

Observe that $\min\{|f(z)| : z \in \overline{D}\}$ may be attained at an interior point of D without the analytic function f on $\overline{D}$ being constant. For example, consider $f(z) = z$, for $|z| < 1$. Then, for $|z| \leq r$ $(r < 1)$,

$$|f(z)| = |z| \geq 0 = |f(0)|$$

so that the minimum modulus of $f(z)$ is attained at the interior point $z = 0$. However, the maximum of $|f(z)|$ on $|z| \leq r$ is attained at $z = r$ which is a boundary point of $|z| < r$.

The minimum principle for harmonic functions is actually stronger than the minimum modulus theorem for analytic functions. The hypothesis that the function be nonzero in the domain is unnecessary for harmonic functions. Of course, a harmonic function can assume negative values in a domain, whereas the modulus of an analytic function cannot.

Corollary 10.10. *Suppose $u(z)$ is harmonic in a bounded domain D whose boundary is the closed contour C. If $u(z)$ is continuous in $D \cup C$, with $u(z) \equiv K (K$ a constant$)$ on C, then $u(z) \equiv K$ throughout D.*

Proof. Since $D \cup C$ forms a compact set, $u(z)$ must attain a maximum and minimum. By Theorem 10.9, the maximum and minimum cannot occur in D. Thus, they must occur on C. But this means that $\max u(z) = \min u(z) = K$. Hence, $u(z) \equiv K$ throughout D. ∎

The boundedness of D in Corollary 10.10 is essential. The domain $\{z : \operatorname{Re} z > 0\}$ has the boundary $\{z : \operatorname{Re} z = 0\}$. The function $u(z) = x$ is continuous for $\operatorname{Re} z \geq 0$ with $u(z) \equiv 0$ on the boundary. But $u(z) \neq 0$ for $\operatorname{Re} z > 0$.

Corollary 10.11. *Suppose $u_1(z)$ and $u_2(z)$ are harmonic in a bounded domain D whose boundary is the closed contour C. If $u_1(z)$ and $u_2(z)$ are continuous in $D \cup C$, with $u_1(z) \equiv u_2(z)$ on C, then $u_1(z) \equiv u_2(z)$ throughout D.*

Proof. Set $u(z) = u_1(z) - u_2(z)$ and apply Corollary 10.10. ∎

Example 10.12. Suppose that $f(z)$ is an entire function such that $f(z)$ is real on the unit circle $|z| = 1$. Then $f(z)$ is constant.

To see this, we set $f = u + iv$. By assumption, $v(z) = 0$ on $|z| = 1$. By Corollary 10.10, $v(z) = 0$ for $|z| < 1$. Hence, $f(z)$ is real for $|z| < 1$, i.e., $f(|z| < 1) \subseteq \mathbb{R}$. By the open mapping theorem, f must be constant for $|z| < 1$. By the uniqueness theorem for analytic functions, f must be a constant throughout $\mathbb{C}$. ∎

There is an interesting relationship between the maximum modulus of an analytic function and the maximum of its real part.

Theorem 10.13. (Borel–Carathéodory) *Suppose $f(z)$ is analytic in the disk $|z| \leq R$. Let $M(r) = \max_{|z|=r} |f(z)|$ and $A(r) = \max_{|z|=r} \operatorname{Re} f(z)$. Then for $0 < r < R$,*

$$M(r) \leq \frac{2r}{R - r} A(R) + \frac{R + r}{R - r} |f(0)|.$$

Proof. If $f(z)$ is constant (say $f(z) = k$), then the right-hand side is bounded below by

$$\frac{-2r}{R-r}|k| + \frac{R+r}{R-r}|k| = |k| = M(r),$$

and the result follows. Hence, we may assume that $f(z)$ is nonconstant.

If $f(0) = 0$, then by Theorem 10.9, $A(R) > A(0) = 0$. Since

$$\operatorname{Re}\{2A(R) - f(z)\} \geq A(R) > 0$$

for $|z| \leq R$, and

$$|2A(R) - f(z)|^2 \geq |f(z)|^2 + 4A(R)[A(R) - \operatorname{Re} f(z)] \geq |f(z)|^2,$$

the function

$$g(z) = \frac{f(z)}{2A(R) - f(z)}$$

is analytic and $|g(z)| \leq 1$ for $|z| \leq R$. Then by Schwarz's lemma,

$$\max_{|z|=r} |g(z)| \leq r/R.$$

But

$$|f(z)| = \left|\frac{2A(R)g(z)}{1+g(z)}\right| \leq \frac{2A(R)r/R}{1-r/R} = \frac{2rA(R)}{R-r}, \tag{10.3}$$

and the result follows when $f(0) = 0$.

Finally, if $f(0) \neq 0$, we apply (10.3) to $f(z) - f(0)$. This leads to

$$|f(z) - f(0)| \leq \frac{2r}{R-r} \max_{|z|=r} \operatorname{Re}\{f(z) - f(0)\} \leq \frac{2r}{R-r}(A(R) + |f(0)|).$$

Thus

$$|f(z)| \leq \frac{2r}{R-r}(A(R) + |f(0)|) + |f(0)| = \frac{2r}{R-r}A(R) + \frac{R+r}{R-r}|f(0)|,$$

and the theorem is proved. ■

Theorem 10.13 may be used to generalize both Theorem 8.35 and Theorem 10.4 as follows.

Theorem 10.14. *Suppose $f(z)$ is an entire function and that $\operatorname{Re} f(z) \leq Mr^\lambda$ for $|z| = r \geq r_0$ and for some nonnegative real number λ. Then $f(z)$ is a polynomial of degree at most $[\lambda]$.*

Proof. Set $R = 2r$ in Theorem 10.13. Then

$$|f(z)| \leq \frac{2r}{2r-r}A(2r) + \frac{2r+r}{2r-r}|f(0)| \leq 2(2r)^\lambda M + 3|f(0)| \leq M_1 r^\lambda,$$

for M_1 sufficiently large. The result now follows from Theorem 8.35. ■

Questions 10.15.

1. When can we say $\ln|f(z)|$ is harmonic? Where is it harmonic?
2. In the proof of Theorem 10.1, where did we use the fact that the domain was simply connected?
3. What theorems are valid for disks but not for a simply connected domain?
4. Where was continuity of the second partial derivatives for harmonic functions important?
5. Can a nonconstant function harmonic in the plane omit more than one real value?
6. Let $f = u + iv$ be analytic in a domain D. Is u_{xx} harmonic in D?
7. Can the maximum modulus theorem for analytic functions be proved using the maximum principle for harmonic functions?
8. Suppose a function is harmonic in a domain D and continuous on its boundary C. Must the function be continuous in $D \cup C$?
9. For a harmonic function u in a domain D which vanishes in an open subset of D, does u vanish identically in D?
10. Is there a relationship between the coefficients of an analytic function and the maximum of its real part?
11. Why is Theorem 10.14 a generalization of Theorem 8.35 and Theorem 10.4?
12. Is every harmonic function an open mapping?
13. Let Ω be a domain and $u \in C^3(\Omega)$. If u is harmonic on Ω, must u_x be harmonic on Ω? Must u_y be harmonic on Ω?
    Note: $C^k(\Omega)$ denotes the set of all functions u whose partial derivatives of order k all exist and are continuous on Ω.
14. What is the average value of the harmonic function $u(x, y) = xy$ on the circle $(x - 2)^2 + (y + 1)^2 = 1$?
15. Let $u(z)$ be harmonic on the disk $|z| < r$ such that $u_x(z) = 0$ on $|z| < r$. What can we conclude about u?
16. Let u be harmonic for $|z| < 1$. Suppose that $\{z_n\}_{n\geq 1}$ is a sequence of complex numbers not equal to z_0 such that $z_n \to z_0$ in $|z| < 1$ and $u(z_n) = 0$ for $n \in \mathbb{N}$. Must u be identically zero? If not, under what additional assumption, do we get $u \equiv 0$?
17. Must a product of two harmonic functions u and v be harmonic?
18. Suppose that u is harmonic in a domain D and v is its harmonic conjugate. Must uv be harmonic on D? Must u^2 be harmonic on D?
19. We know that $u(z) = \ln|z|$ is harmonic in the annulus $D = \{z : 1 < |z| < 2\}$. Can $u(z)$ have a harmonic conjugate on D?

Exercises 10.16.

1. Show that a function harmonic in a domain must have partial derivatives of all orders.

2. If $u(z)$ is nonconstant and harmonic in the plane, show that $u(z)$ comes arbitrarily close to every real value.
3. Prove the minimum principle directly by each of the three methods in which the maximum principle was proved.
4. Show that $\int_0^\pi \ln \sin\theta \, d\theta = -\pi \ln 2$ by applying the mean-value principle to $\ln|1+z|$ for $|z| \le r < 1$, and then letting $r \to 1$.
5. Suppose $f(z)$ and $g(z)$ are analytic inside and on a simple closed contour C, with $\operatorname{Re} f(z) = \operatorname{Re} g(z)$ on C. Show that $f(z) = g(z) + i\beta$ inside C, where β is a real constant.
6. Generalize the previous exercise by showing that the conclusion still holds if it is only assumed that $f(z)$ and $g(z)$ are analytic inside C and continuous in the region consisting of C and its interior.
7. If $u(z)$ is harmonic and bounded in the punctured disk $0 < |z-z_0| < R$, show that $\lim_{z\to z_0} u(z)$ exists.
8. Suppose $u_1(z)$ and $u_2(z)$ are harmonic in a simply connected domain D, with $u_1(z)u_2(z) \equiv 0$ in D. Prove that either $u_1(z) \equiv 0$ or $u_2(z) \equiv 0$ in D.
9. It is easy to see that $u(z) = \operatorname{Im}\left(\frac{1+z}{1-z}\right)^2$ is harmonic in the unit disk $|z| < 1$ and $\lim_{r\to 1^-} u(re^{i\theta}) = 0$ for all θ. Why does this not contradict the maximum principle for harmonic functions? Is u continuous on $|z| = 1$?
10. Does there exist a harmonic function in $|z| < 1$ taking the value 1 everywhere on $|z| = 1$? Is your solution unique?
11. Does there exist a harmonic function on the strip $\{z : 0 < \operatorname{Re} z < 1\}$ with $u(x,0) = 0$ and $u(x,1) = 1$? Is your solution unique?
12. If $u(z) = u(x,y)$ is harmonic in the plane with $u(z) \le |z|^n$ for every z, show that $u(z)$ is a polynomial in the two variables x and y.
13. Suppose that $f(z)$ is analytic in the disk $|z| \le R$, and let $A(r) = \max_{|z|=r} \operatorname{Re} f(z)$. Prove that for $r < R$,
$$\max_{|z|=r} \frac{|f^{(n)}(z)|}{n!} \le \frac{2^{n+2}R}{(R-r)^{n+1}}\{A(r) + |f(0)|\}.$$

10.2 Poisson Integral Formula

In this section, we shall attempt to find a harmonic analog to Cauchy's integral formula. If f is analytic inside and on a simple closed contour C, then

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - z}\, d\zeta \tag{10.4}$$

at all points z inside C. We would like to find an expression for $\operatorname{Re} f$ at points inside C in terms of the values of $\operatorname{Re} f$ on C. Unfortunately, the expression

$$\mathrm{Re}\left\{\frac{1}{2\pi i}\int_C \frac{f(\zeta)}{\zeta - z}\,d\zeta\right\}$$

simplifies into one involving both $\mathrm{Re}\, f$ and $\mathrm{Im}\, f$ on C.

If, however, the integral of (10.4) is transformed into one of the form $\int_a^b \phi(t)\,dt$, where $\phi(t)$ is a complex-valued function of a real variable t, then

$$\mathrm{Re}\int_a^b \phi(t)\,dt = \int_a^b \mathrm{Re}\,\phi(t)\,dt.$$

Recall that we performed this kind of transformation when proving the mean-value principle for harmonic functions. This enabled us to determine the value of a harmonic function at the center of a circle based on its values on the circumference. By (10.4), we have the so-called mean value property for analytic functions:

$$f(a) = \frac{1}{2\pi}\int_0^{2\pi} f(a + re^{i\phi})\,d\phi, \quad 0 < r < \mathrm{dist}\,(a, C) = R$$

for a inside C. The value $f(a)$ of f at the center a of the disk $|z - a| < r$ is expressed by the integration of f over the boundary circle $|z - a| = r$ of this disk. Note that $f(a)$ is the same for all r in the interval $(0, R)$. We wish to obtain similar expression for a point of the disk $|z - a| < r$ other than the center. But an analog to the Cauchy integral formula for the circle is an expression for the harmonic function at all points inside the circle in terms of its values on the circle.

Lemma 10.17. (Poisson Integral Formula for Analytic Functions) *Suppose $f(z)$ is analytic in a domain containing the closed unit disk $|z| \le 1$. Then for $|z| < 1$, we have*

$$f(z) = \frac{1}{2\pi}\int_{|\zeta|=1} \frac{1 - |z|^2}{|\zeta - z|^2} f(\zeta)\,\frac{d\zeta}{i\zeta}, \tag{10.5}$$

or equivalently,

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1 - |z|^2}{|e^{i\phi} - z|^2} f(e^{i\phi})\,d\phi. \tag{10.6}$$

Proof. By Cauchy's integral formula, we have

$$f(z) = \frac{1}{2\pi i}\int_{|\zeta|=1} \frac{f(\zeta)}{\zeta - z}\,d\zeta = \frac{1}{2\pi}\int_0^{2\pi} \frac{\zeta f(\zeta)}{\zeta - z}\,d\phi \quad (|z| < 1). \tag{10.7}$$

If $z = 0$, the result follows from Gauss's mean-value theorem. So we may suppose that $z \neq 0$, and set

$$z^* = 1/\overline{z}$$

which is the reflection of z in the unit circle. The point z^*, which lies on the ray from the origin through z, is outside the unit circle $|\zeta| = 1$. Hence (as $z^* = 1/\overline{z}$), for $|z| < 1$

$$0 = \frac{1}{2\pi i} \int_{|\zeta|=1} \frac{f(\zeta)}{\zeta - z^*} \, d\zeta = -\frac{1}{2\pi i} \int_{|\zeta|=1} \frac{\overline{z} f(\zeta)}{1 - \zeta \overline{z}} \, d\zeta. \tag{10.8}$$

Subtracting (10.8) from (10.7), we get

$$f(z) = \frac{1}{2\pi i} \int_{|\zeta|=1} \left[\frac{1}{\zeta - z} + \frac{\overline{z}}{1 - \zeta \overline{z}} \right] f(\zeta) \, d\zeta. \tag{10.9}$$

We can simplify (since $|\zeta| = 1$) to

$$\frac{1}{\zeta - z} + \frac{\overline{z}}{1 - \zeta \overline{z}} = \frac{1 - |z|^2}{(\zeta - z)(1 - \zeta \overline{z})} = \frac{1 - |z|^2}{(\zeta - z)(\overline{\zeta} - \overline{z})\zeta} = \frac{1 - |z|^2}{|\zeta - z|^2} \frac{1}{\zeta}.$$

Using the last equality, (10.9) gives (10.5). Equation (10.6) follows if we let $\zeta = e^{i\theta}$ in (10.5). ∎

The following general result is a consequence of Lemma 10.17.

Theorem 10.18. (Poisson Integral Formula for analytic functions) *Suppose $f(z)$ is analytic in a domain containing the closed disk $|z - a| \leq R$. Then for $|z - a| < R$, we have*

$$f(z) = \frac{1}{2\pi} \int_{|\zeta - a| = R} \frac{R^2 - |z - a|^2}{|\zeta - z|^2} f(\zeta) \, \frac{d\zeta}{i(\zeta - a)},$$

or equivalently,

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - |z - a|^2}{|Re^{i\phi} - (z - a)|^2} f(a + Re^{i\phi}) \, d\phi. \tag{10.10}$$

Proof. By the change of variable $w = (z - a)/R$, it reduces to the case where $R = 1$ and $a = 0$. ∎

In particular, for $a = 0$, the formula reduces to

$$f(re^{i\theta}) = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - r^2}{|Re^{i\phi} - re^{i\theta}|^2} f(Re^{i\phi}) \, d\phi.$$

The expression (with $\zeta = Re^{i\phi}$, $z = re^{i\theta}$ and $r < R$)

$$P(z, \zeta) = \frac{|\zeta|^2 - |z|^2}{|\zeta - z|^2} = \mathrm{Re} \left(\frac{\zeta + z}{\zeta - z} \right) = \frac{R^2 - r^2}{R^2 - 2rR\cos(\theta - \phi) + r^2}$$

is known as the *Poisson kernel* for the disk $|z| < R$. Note that the Poisson kernel is bounded above by

$$\frac{R^2 - r^2}{R^2 - 2rR + r^2} = \frac{R+r}{R-r},$$

and is bounded below by

$$\frac{R^2 - r^2}{R^2 + 2rR + r^2} = \frac{R-r}{R+r}.$$

Let $a = 0$ and let $f(z) = u(z) + iv(z)$ be analytic for $|z| \leq R$. Then, from Theorem 10.18, it follows that

$$f(z) = \frac{1}{2\pi} \int_{|\zeta|=R} P(z,\zeta) f(\zeta)\, d\phi$$

and, equating the real part gives

Theorem 10.19. (Poisson Integral Formula for Harmonic Functions) *Suppose $u(z)$ is harmonic in a domain containing the disk $|z| \leq R$. Then for $z = re^{i\theta}$, $r < R$, we have*

$$u(z) = \frac{1}{2\pi} \int_{|\zeta|=R} P(z,\zeta) u(\zeta)\, d\phi;$$

or equivalently,

$$u(re^{i\theta}) = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - r^2}{R^2 - 2rR\cos(\theta - \phi) + r^2} u(Re^{i\phi})\, d\phi.$$

A similar formula holds for the imaginary part $v(z)$ of $f(z)$.

Corollary 10.20. *For $r < R$ and θ arbitrary,*

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - r^2}{R^2 - 2rR\cos(\theta - \phi) + r^2}\, d\phi = \frac{1}{2\pi} \int_0^{2\pi} P(re^{i\theta}, Re^{i\phi})\, d\phi = 1.$$

Proof. Set $u(z) \equiv 1$ in Theorem 10.19. ∎

Theorem 10.21. *Suppose $f(z) = u(z) + iv(z)$ is analytic in the disk $|z| \leq 1$. Then for $|z| < 1$, we may express $f(z)$ as*

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\zeta + z}{\zeta - z} u(\zeta)\, d\phi + iv(0) \quad (\zeta = e^{i\phi}). \tag{10.11}$$

Proof. To do this it suffices to recall (10.7) and (10.8):

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\zeta}{\zeta - z} f(\zeta)\, d\phi, \tag{10.12}$$

and (because $\zeta\overline{\zeta} = 1$),

$$0 = -\frac{1}{2\pi i}\int_{|\zeta|=1} \frac{\overline{z}}{1-\zeta\overline{z}} f(\zeta)\, d\zeta = -\frac{1}{2\pi}\int_0^{2\pi} \frac{\overline{z}}{\overline{\zeta}-\overline{z}} f(\zeta)\, d\phi.$$

Since the integral on the right is a Riemann integral, taking conjugation on the right leads to

$$0 = -\frac{1}{2\pi}\int_0^{2\pi} \frac{z}{\zeta - z}\overline{f(\zeta)}\, d\phi. \tag{10.13}$$

Writing $f(\zeta) = u(\zeta) + iv(\zeta)$, and then adding (10.12) and (10.13) shows that

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\zeta + z}{\zeta - z} u(\zeta)\, d\phi + i\frac{1}{2\pi}\int_0^{2\pi} v(\zeta)\, d\phi.$$

The desired formula (10.11) follows if we apply the mean-value property for the last integral to the harmonic function v. ∎

Equation (10.11) (also known as the Schwarz formula) determines the analytic function $f(z)$ within an additive imaginary constant once its real part on the unit circle is given. Thus, Schwarz formula is considered to be more powerful than the Poisson integral formula. Nevertheless, the latter is a fundamental formula in mathematical physics and fluid mechanics. More generally, by the change of variable $w = z/R$, Theorem 10.21 gives

Theorem 10.22. *Suppose $f(z) = u(z) + iv(z)$ is analytic in the disk $|z| \leq R$. Then for $|z| < R$, we may express $f(z)$ as*

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\zeta + z}{\zeta - z} u(\zeta)\, d\phi + iv(0) \quad (\zeta = Re^{i\phi}); \tag{10.14}$$

or equivalently,

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{Re^{i\phi} + re^{i\theta}}{Re^{i\phi} - re^{i\theta}} u(Re^{i\phi})\, d\phi + iv(0).$$

(*The integral on the right is called the complex Poisson integral*).

Equating imaginary part on both sides of (10.14) gives

Corollary 10.23. *Suppose $f(z) = u(z) + iv(z)$ is analytic in the disk $|z| \leq R$. Then for $\zeta = Re^{i\phi}$, $z = re^{i\theta}$ and $r < R$, we may also express $v(z)$ as*

$$v(z) = \frac{1}{2\pi}\int_{|\zeta|=R} \mathrm{Im}\left(\frac{\zeta + z}{\zeta - z}\right) u(\zeta)\, d\phi + v(0),$$

or equivalently

$$v(re^{i\theta}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{2rR\sin(\theta - \phi)}{R^2 - 2rR\cos(\theta - \phi) + r^2} u(Re^{i\phi})\, d\phi + v(0).$$

Remark 10.24. We know in general (Exercise 10.16(5)) that an analytic function is determined to within an imaginary constant by its real part. In the case that the function $f(z) = u(z) + iv(z)$ is analytic in the disk $|z| \le R$, Theorem 10.22 gives this relationship explicitly. ●

As we shall now see, the conclusion of Theorem 10.19 is valid under less stringent conditions. It need not be assumed that $u(z)$ is harmonic on the circle $|z| = R$. The proof requires an acquaintance with the notion of uniform convergence. Again it suffices to deal with $R = 1$, since the general case follows from a simple transformation.

Theorem 10.25. *Suppose $u(z)$ is harmonic in the open disk $|z| < 1$ and continuous on the closed disk $|z| \le 1$. Then for $z = re^{i\theta}$, $r < 1$, we have*

$$u(re^{i\theta}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-r^2}{1-2r\cos(\theta-\phi)+r^2}\, u(e^{i\phi})\, d\phi.$$

Proof. Let $f(z) = u(z) + iv(z)$ be analytic for $|z| < 1$, and let $\{t_n\}$ be an increasing sequence of positive real numbers approaching 1. Then for each n, define

$$f_n(z) = f(t_n z), \quad u_n(z) = u(t_n z), \quad \text{and} \quad v_n(z) = v(t_n z).$$

Clearly, $v_n(0) = v(0)$ for each n and

$$u_n(z) = \operatorname{Re} f(t_n z), \quad \text{and} \quad v_n(z) = \operatorname{Im} f(t_n z).$$

As $u(t_n z)$ is harmonic in the closed disk $|z| \le 1$, we obtain that $f(t_n z)$ is analytic in the closed disk $|z| \le 1$ (since $1/t_n > 1$), and so Theorem 10.21 is applicable for f_n. Thus, for each fixed z with $|z| < 1$,

$$f_n(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{e^{i\phi}+z}{e^{i\phi}-z} u_n(e^{i\phi})\, d\phi + iv_n(0).$$

Since $f_n(z)$ is continuous at z ($|z| < 1$) and $t_n z \to z$ as $n \to \infty$,

$$\lim_{n\to\infty} f_n(z) = \lim_{n\to\infty} f(t_n z) = f(z), \quad |z| < 1.$$

The proof will be completed by verifying that

$$\int_0^{2\pi} \frac{e^{i\phi}+z}{e^{i\phi}-z} u_n(e^{i\phi})\, d\phi \to \int_0^{2\pi} \frac{e^{i\phi}+z}{e^{i\phi}-z} u(e^{i\phi})\, d\phi. \tag{10.15}$$

(Recall that $v_n(0) = v(0)$). It suffices to show that the difference

$$\left|\int_0^{2\pi} \frac{e^{i\phi}+z}{e^{i\phi}-z} (u_n(e^{i\phi}) - u(e^{i\phi}))\, d\phi\right| \le \frac{1+r}{1-r}\int_0^{2\pi} \left|u_n(e^{i\phi}) - u(e^{i\phi})\right|\, d\phi$$

can be made arbitrarily small. Note that, $u(z)$, being continuous on the compact set $|z| \le 1$, is uniformly continuous on $|z| \le 1$. So

$$u_n(e^{i\phi}) = u(t_n e^{i\phi}) \to u(e^{i\phi})$$

uniformly with respect ϕ, $0 \le \phi \le 2\pi$. Consequently, the expression on the last integral converges to zero as $n \to \infty$. Thus, for $|z| < 1$, we have

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\phi} + z}{e^{i\phi} - z} u(e^{i\phi})\, d\phi + iv(0).$$

Equating the real part on both sides, we have the desired result. ■

Remark 10.26. The uniform continuity of $u(z)$ $(|z| \le 1)$ enabled us to show that the sequence $u_n(z) = u(t_n z)$ converged uniformly to $u(z)$ $(|z| \le 1)$. Thus, the validity of (10.15) is a consequence of Theorem 8.11. ●

By a simple transformation, Theorem 10.25 shows that a function, harmonic for $|z| < R$ and continuous for $|z| \le R$, has the property that its values inside the disk are determined by its values on the boundary. Suppose, instead, that we start with a real-valued function $F(\theta)$ continuous on the circle $|z| = R$. Does there exist a function $u(z)$ harmonic in the disk $|z| < R$ having prescribed boundary values? More generally, the *Dirichlet problem* deals with the following question: Given a domain D, and a function $F : \partial D \to \mathbb{R}$, does there exist a function u that is harmonic in D such that $u = F$ on the boundary ∂D? The solution to this problem has immediate applications in fluid mechanics. Our next theorem solves the Dirichlet problem for the disk.

Theorem 10.27. (Schwarz's Theorem) *Let F be a continuous function of a real variable defined on the unit circle $|\zeta| = 1$. Then the real-valued function $u(z)$ defined by*

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} P(z, e^{i\phi}) F(e^{i\phi})\, d\phi \quad (|z| < 1)$$

is harmonic in the disk $|z| < 1$, and for each fixed t, $0 \le t \le 2\pi$,

$$\lim_{z \to e^{it}} u(z) := \lim_{\substack{r \to 1- \\ \theta \to t}} u(re^{i\theta}) = F(e^{it}) \quad (|z| < 1).$$

(*In addition, if we let $u(z) = F(z)$ for $|z| = 1$, then $u(z)$ becomes continuous for $|z| \le 1$*).

Proof. First we verify that the function u defined in the statement is harmonic in the disk $|z| < 1$. To see this, we may rewrite

$$\begin{aligned} u(z) &= \mathrm{Re}\left[\frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\phi} + z}{e^{i\phi} - z} F(e^{i\phi})\, d\phi\right] \\ &= \mathrm{Re}\left[\frac{1}{2\pi} \int_{|\zeta|=1} \frac{\zeta + z}{\zeta - z} F(\zeta)\, \frac{d\zeta}{i\zeta}\right] \end{aligned}$$

$$= \mathrm{Re}\left[\frac{1}{2\pi i}\int_{|\zeta|=1}\left(\frac{2}{\zeta - z} - \frac{1}{\zeta}\right)F(\zeta)\,d\zeta\right]$$
$$= \mathrm{Re}\,[2f(z) - f(0)]\,, \quad f(z) = \frac{1}{2\pi i}\int_{|\zeta|=1}\frac{F(\zeta)}{\zeta - z}\,d\zeta.$$

Note that $f(z)$ is an analytic function in $|z| < 1$, and

$$f(0) = \frac{1}{2\pi i}\int_{|\zeta|=1}\frac{F(\zeta)}{\zeta}\,d\zeta = \frac{1}{2\pi}\int_0^{2\pi}F(e^{i\phi})\,d\phi = F(0) \in \mathbb{R}.$$

Thus, $u(z)$ is a real part of an analytic function $2f(z) - f(0)$ for $|z| < 1$, and therefore it is harmonic in $|z| < 1$.

To prove the second part, we must show that to each fixed ζ ($|\zeta| = 1$),

$$u(z) \to F(\zeta) \text{ as } z \to \zeta := e^{it} \quad (|z| < 1),$$

on the assumption that F is a continuous function of t, $0 \le t \le 2\pi$. Thus, we need to show that, for each $\epsilon > 0$, there corresponds a $\delta > 0$ such that

$$\left|u(z) - F(e^{it})\right| < \epsilon$$

for all z satisfying $|z - e^{it}| < \delta$. To do this, we first recall that

$$1 = \frac{1}{2\pi}\int_0^{2\pi}P(z, e^{i\phi})\,d\phi \quad (|z| < 1),$$

and noting that $P(z, e^{i\phi}) > 0$, we consider the expression

$$u(z) - F(e^{it}) = \frac{1}{2\pi}\int_0^{2\pi}P(z, e^{i\phi})[F(e^{i\phi}) - F(e^{it})]\,d\phi. \tag{10.16}$$

From the definition of the Poisson kernel

$$P(re^{it}, e^{it}) = \mathrm{Re}\left(\frac{e^{it} + re^{it}}{e^{it} - re^{it}}\right) = \frac{1+r}{1-r}$$

so that $\lim_{r\to 1} P(re^{it}, e^{it}) = \infty$, whereas for $\theta \neq t$ with $|\theta - t| < \pi$,

$$\lim_{r\to 1} P(re^{i\theta}, e^{it}) = \lim_{r\to 1}\mathrm{Re}\left(\frac{e^{it} + re^{i\theta}}{e^{it} - re^{i\theta}}\right) = \mathrm{Re}\left(\frac{e^{it} + e^{i\theta}}{e^{it} - e^{i\theta}}\right) = 0.$$

To show that (10.16) can be made arbitrarily small in absolute value, we have for a small $\delta > 0$

$$|u(z) - F(e^{it})| \le \frac{1}{2\pi}\int_{t-\delta}^{t+\delta}P(z, e^{i\phi})|F(e^{i\phi}) - F(e^{it})|\,d\phi$$
$$+ \frac{1}{2\pi}\int_{t+\delta}^{2\pi+t-\delta}P(z, e^{i\phi})|F(e^{i\phi}) - F(e^{it})|\,d\phi.$$

Note that the integrand in (10.16) is periodic in ϕ of period 2π and so, we have used the fact that $\int_0^{2\pi} = \int_{t-\delta}^{2\pi+t-\delta} = \int_{t-\delta}^{t+\delta} + \int_{t+\delta}^{2\pi+t-\delta}$. By the continuity of F at t, for an arbitrary $\epsilon > 0$, there is a small $\delta > 0$ such that

$$|F(e^{i\phi}) - F(e^{it})| < \epsilon \quad \text{whenever } |\phi - t| < \delta.$$

We use this for the first integral in the last inequality. On the other hand, for $|\phi - t| \geq \delta$ and $|\arg z - \arg \zeta| = |\arg z - t| < \delta/2$ with $z = re^{i\theta}$, we have (see Figure 10.1)

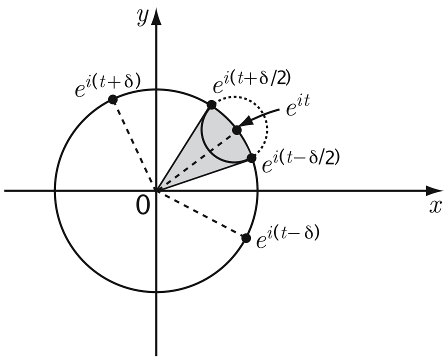


Figure 10.1.

$$|\theta - \phi| = |\theta - t - (\phi - t)| \geq |\phi - t| - |\theta - t| = \delta - \delta/2 = \delta/2.$$

Therefore, for $|\phi - t| \geq \delta$ and $|\theta - t| < \delta/2$, we have

$$\begin{aligned} P(z, e^{i\phi}) &= \frac{1 - r^2}{1 - 2r\cos(\theta - \phi) + r^2} \\ &= \frac{1 - r^2}{(1 - r)^2 + 2r(1 - \cos(\theta - \phi))} \\ &= \frac{1 - r^2}{(1 - r)^2 + 4r\sin^2((\theta - \phi)/2)} \\ &< \frac{1 - r^2}{4r\sin^2(\delta/4)} \end{aligned}$$

and use this for the second integral. Thus, for $|\arg z - t| < \delta/2$,

$$\begin{aligned} |u(z) - F(e^{it})| &\leq \frac{\epsilon}{2\pi} \int_{t-\delta}^{t+\delta} P(z, e^{i\phi})\, d\phi \\ &\quad + \frac{1}{2\pi}\left[2 \max_{|\zeta|=1} |F(\zeta)| \frac{1 - r^2}{4r\sin^2(\delta/4)} (2\pi - 2\delta)\right] \\ &< \epsilon + \max_{|\zeta|=1} |F(\zeta)| \frac{1 - r^2}{2r\sin^2(\delta/4)}. \end{aligned}$$

The second term on the right can be made less than any $\epsilon > 0$ for r close to 1. Thus, there exists a $\delta' > 0$ such that

$$\left|u(z) - F(e^{it})\right| < 2\epsilon \;\text{ whenever } z \text{ with } |z| < 1 \text{ and } |z - e^{it}| < \delta'.$$

Thus, $\lim_{z \to e^{it}} u(z) = F(e^{it})$. ■

Remark 10.28. A slight modification in the above proof shows that any function satisfying the conditions of the theorem must be continuous on the closed disk $|z| \le 1$. In view of Corollary 10.11, the function $u(z)$ of Theorem 10.19 (with $R = 1$) is the only function that can satisfy the conditions of the theorem. Using the Riemann mapping theorem which will be proved in Chapter 11, we see that Dirichlet's problem can be solved for simply connected domain D. ●

A simple translation applied to Theorem 10.27 leads to a general result which we formulate as follows.

Theorem 10.29. *Let $F(\phi) := F(Re^{i\phi})$ be a continuous function of the real variable ϕ, $0 \le \phi \le 2\pi$, with $F(0) = F(2\pi)$. Then the function $u(z)$ defined by*

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} P(z, Re^{i\phi}) F(Re^{i\phi})\, d\phi \quad (|z| < R)$$

satisfies the following conditions:

(i) *$u(z)$ is harmonic in the disk $|z| < R$.*
(ii) *For each fixed t, $0 \le t \le 2\pi$,*

$$\lim_{z \to Re^{it}} u(z) = F(Re^{it}) \quad (|z| < R).$$

Remark 10.30. By requiring in Theorem 10.29 only that the function $F(\phi)$ be sectionally continuous, the conclusion (i) still holds with the restriction that $\lim_{r \to R} u(re^{i\theta}) = F(\theta)$ only at the points of continuity for F. The proof is identical. Again, the analog of Theorem 10.27 for $|z - z_0| \le R$ follows routinely as well. ●

Example 10.31. Suppose that we wish to find a real-valued function u harmonic in the open first quadrant $D = \{z = x + iy : x, y > 0\}$, continuous on $\overline{D} \setminus \{0\}$ and $u(x, 0) = 5$ for $x > 0$ and $u(0, y) = 3$ for $y > 0$.

To do this, we may consider

$$\operatorname{Log} z = \ln |z| + i \operatorname{Arg} z \;\text{ on }\; D_\pi = \mathbb{C} \setminus (-\infty, 0].$$

Then $v(x, y) = \operatorname{Arg} z$ is harmonic on D_π,

$$v(x, 0) = 0 \text{ for } x > 0 \;\text{ and }\; v(0, y) = \pi/2 \text{ for } y > 0.$$

To obtain u satisfying the desired properties, we define

$$u(x,y) = av(x,y) + b.$$

To find a and b, we set $y = 0$ and obtain

$$5 = u(x,0) = av(x,0) + b = a(0) + b, \quad \text{i.e.,} \quad b = 5.$$

Setting $x = 0$, we have

$$3 = u(0,y) = av(0,y) + b = a(\pi/2) + b, \quad \text{i.e.,} \quad a = -4/\pi.$$

The desired function is then

$$u(x,y) = -(4/\pi)\operatorname{Arg} z + 5.$$

This problem can be also solved by using the Poisson integral formula for the half-plane (see Exercise 10.35(10)). ■

Remark 10.32. As remarked before, the results of this section that are stated for unit disks could be stated for arbitrary disks. To illustrate, suppose $u(z)$ is harmonic in a domain containing the disk $|z - z_0| \leq R$. Setting $z - z_0 = re^{i\theta}$, the conclusion of Theorem 10.19 remains valid for any point inside the circle $|z - z_0| = R$. ●

Example 10.33. Solve the Dirichlet problem:

$$u_{xx} + u_{yy} = 0, \quad -\infty < x < \infty, \ y > 0,$$

subject to $u(x,0) = 0$ for $|x| > 1$ and $u(x,0) = x$ for $x \in (-1,1)$, see Figure 10.2. According to Exercise 10.35(10)

$$u(x,y) = \frac{y}{\pi}\int_{-1}^{1} \frac{t\,dt}{(x-t)^2 + t^2}$$

and a simple computation gives

$$u(x,y) = \frac{x}{\pi}\left[\tan^{-1}\frac{x+1}{y} - \tan^{-1}\frac{x-1}{y}\right] + \frac{y}{2\pi}\ln\frac{(x-1)^2+y^2}{(x+1)^2+y^2}. \quad ●$$

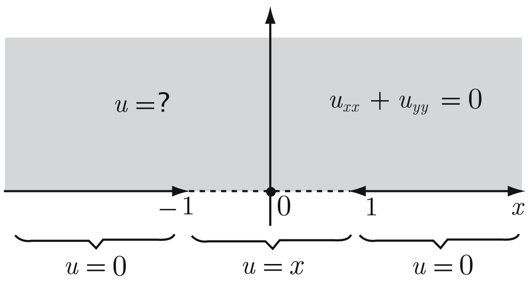


Figure 10.2.

Questions 10.34.

1. What properties of the point z^*, chosen in the proof of Theorem 10.19, made the proof work?
2. Can Theorem 10.14 be proved using Poisson's formula?
3. If $F(\phi)$, defined on the circle $|z| = R$, is continuous at all but a finite number of points, is there a unique function u harmonic for $|z| < R$ that approaches $F(\phi)$ as z approaches the boundary?
4. Can the solution to the Dirichlet problem for the disk be used to solve a Dirichlet problem for different regions?
5. What is the relationship between Theorem 10.6 and Theorem 10.19?

Exercises 10.35.

1. If $f(z)$ is a continuous function on $|z| = 1$, show that $F(z)$ defined by
$$F(z) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{f(\zeta)}{\zeta - z}\, d\zeta$$
is analytic for $|z| < 1$.
2. (a) For $\rho = Re^{i\phi}$, $z = re^{i\theta}$ $(r < R)$, show that
$$\frac{\rho + z}{\rho - z} = 1 + 2\sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^n e^{in(\theta-\phi)}.$$
(b) Conclude that
$$\frac{R^2 - r^2}{R^2 - 2rR\cos\alpha + r^2} = 1 + 2\sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^n \cos n\alpha \quad (\alpha \text{ real}).$$
3. Use the previous exercise to find an alternate expression for the conclusion of Theorem 10.19.
4. Show that
$$\int_0^{2\pi} \frac{\sin(\theta - \phi)}{R^2 - 2rR\cos(\theta - \phi) + r^2}\, d\phi = 0.$$
5. If $u(z)$ is harmonic for $|z| > R$ and continuous for $|z| \geq R$, show that for $\rho = Re^{i\phi}$, $z = re^{i\theta}$ $(r > R)$,
$$u(z) = -\frac{1}{2\pi}\int_0^{2\pi} \operatorname{Re}\frac{\rho + z}{\rho - z} u(Re^{i\theta})\, d\phi.$$
6. Find a function $u(z)$ harmonic in the disk $|z| < R$ for which
$$\lim_{r\to R} u(re^{i\theta}) = \begin{cases} 0 & \text{if } 0 < \theta < \pi \\ 1 & \text{if } \pi < \theta < 2\pi. \end{cases}$$

7. Show that the function
$$u(re^{i\theta}) = \frac{2}{\pi}\tan^{-1}\frac{2r\sin\theta}{1-r^2} \quad (r<1)$$
is harmonic for $|z|<1$ and satisfies the boundary conditions
$$\lim_{r\to 1} u(re^{i\theta}) = \begin{cases} 1 & \text{if } 0<\theta<\pi \\ -1 & \text{if } \pi<\theta<2\pi. \end{cases}$$
8. Set $F(\theta)=\theta/2$, $0\le\theta\le 2\pi$. Show that the function
$$u(re^{ei\theta}) = \tan^{-1}\frac{r\sin\theta}{1+r\cos\theta}$$
is harmonic for $|z|<1$, and that $\lim_{r\to 1} u(re^{i\theta}) = F(\theta)$ for all θ. Can you derive this from the Poisson integral formula?
9. Suppose $f=u+iv$ is analytic and bounded on the real line and the upper half-plane. Show that for $z=x+iy$, $y>0$, we have
$$u(z) = \frac{y}{\pi}\int_{-\infty}^{\infty}\frac{u(t,0)}{(t-x)^2+y^2}\,dt, \quad v(z) = \frac{y}{\pi}\int_{-\infty}^{\infty}\frac{v(t,0)}{(t-x)^2+y^2}\,dt.$$
These are called the Poisson integral formula for u and v in the upper half-plane.
10. Using the previous exercise, formulate and solve a Dirichlet problem for a half-plane. More precisely, prove the following: If $F(x)$ is a continuous function on $\mathbb{R}$, then show that the function $u(x,y)$ defined by
$$u(x,y) = \frac{y}{\pi}\int_{-\infty}^{\infty}\frac{F(t)}{(x-t)^2+y^2}\,dt$$
is a solution of the Dirichlet problem in the upper half-plane $\operatorname{Im} z>0$ with the boundary condition $u(x,0)=F(x)$ for $x\in\mathbb{R}$.
11. Suppose that $f=u+iv$ is analytic for $|z|<1$. Show that for $|z|<r$ $(0<r<1)$,
$$\frac{f^{(n)}(z)}{n!} = \frac{1}{2\pi i}\int_{|\zeta|=r}\frac{2}{(\zeta-z)^{n+1}}u(\zeta)\,d\zeta.$$
12. Find a harmonic function u on the upper half-plane $\operatorname{Im} z>0$ such that $u(x,0)=0$ for $x>0$ and $u(x,0)=1$ for $x<0$.
13. Find a harmonic function u on the upper half-plane $\operatorname{Im} z>0$ such that
    (a) $u(x,0)=1$ for $x<-1$
    (b) $u(x,0)=2$ for $-1<x<1$
    (c) $u(x,0)=3$ for $x>1$.
14. Find a function $u(z)$ harmonic for $|z|<1$ such that
$$\lim_{r\to 1} u(re^{i\theta}) = \begin{cases} 1 & \text{if } 0<\theta<\pi \\ 0 & \text{if } \pi<\theta<2\pi. \end{cases}$$
15. Suppose that $f=u+iv$ is entire and $z^{-1}\operatorname{Re} f(z)\to 0$ as $z\to\infty$. Show that f is a constant.

10.3 Positive Harmonic Functions

As an application of Poisson's integral formula, we prove

Theorem 10.36. (Harnack's Inequality) *Suppose $u(z)$ is harmonic in the disk $\Delta(z_0; R) = \{z : |z - z_0| < R\}$, with $u(z) \geq 0$ for all $z \in \Delta(z_0; R)$. Then for every z in this disk, we have*

$$u(z_0)\frac{R - |z - z_0|}{R + |z - z_0|} \leq u(z) \leq u(z_0)\frac{R + |z - z_0|}{R - |z - z_0|}.$$

Proof. Fix $z \in \Delta(z_0; R)$, and let $s < R$. Then, for every s with $s < R$, the Poisson integral formula given by Theorem 10.18 leads to

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{s^2 - |z - z_0|^2}{|se^{i\phi} - (z - z_0)|^2} u(z_0 + se^{i\phi})\, d\phi \tag{10.17}$$

for every $z \in \Delta(z_0; s)$. Using the positivity of $u(z)$ and the inequality

$$\frac{s - |z - z_0|}{s + |z - z_0|} \leq \frac{s^2 - |z - z_0|^2}{|se^{i\phi} - (z - z_0)|^2} \leq \frac{s + |z - z_0|}{s - |z - z_0|}$$

we get, from (10.17),

$$\frac{s - |z - z_0|}{s + |z - z_0|} u(z_0) \leq u(z) \leq \frac{s + |z - z_0|}{s - |z - z_0|} u(z_0)$$

because, by the mean-value property (i.e., (10.17) for $z = z_0$),

$$u(z_0) = \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + se^{i\phi})\, d\phi.$$

Since the last inequalities are valid whenever $|z - z_0| \leq s < R$, these inequalities continues to hold when s approaches R. ∎

Using Harnack's inequality we can present an alternate proof of Liouville's theorem for harmonic functions (Theorem 10.4) in the following form.

Corollary 10.37. *If u is harmonic in $\mathbb{C}$ and is bounded above (or below), then u is constant.*

Proof. It suffices to prove for $u(z) \geq 0$ in $\mathbb{C}$. Fix z ($|z| = r$) and let $R > r$. By Harnack's inequality

$$\frac{R - r}{R + r} u(0) \leq u(re^{i\theta}) \leq \frac{R + r}{R - r} u(0).$$

Letting $R \to \infty$, we see that $u(z) \leq u(0)$ so that u attains its maximum at $z = 0$ and therefore, u is constant for $|z| < r$ and hence in $\mathbb{C}$. ∎

It was shown (see Theorem 10.6) that every harmonic function satisfies the mean-value property. For continuous functions, the converse is also true.

Theorem 10.38. *Suppose $u(z)$ is a real-valued continuous function such that for each point z_0 in a domain D,*

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{i\theta})\, d\theta$$

whenever the disk $|z - z_0| \leq r$ is contained in D. Then $u(z)$ is harmonic throughout D.

Proof. Choose a point $z_0 \in D$ and $r > 0$ such that $|z - z_0| \leq r$ is contained in D. As a consequence of Theorem 10.29, there exists a function $u_1(z)$ harmonic for $|z - z_0| < r$, continuous for $|z - z_0| \leq r$, and equal to $u(z)$ on the circle $|z - z_0| = r$. Since $u_1(z) - u(z)$ is a continuous function that satisfies the mean-value property, the first proof of Theorem 10.9 shows that $u_1(z) - u(z)$ attains both its maximum and minimum on the boundary. Because

$$u_1(z) - u(z) \equiv 0 \quad \text{on } |z - z_0| = r,$$

it follows that $u_1(z) \equiv u(z)$ for $|z - z_0| < r$. Hence $u(z)$ is harmonic in a neighborhood of z_0. Since z_0 was arbitrary, $u(z)$ is harmonic in D. ∎

Thus, a necessary and sufficient condition for a continuous function to be harmonic in a domain is that it satisfies the mean-value property at each point in the domain. As an application, we prove the following analog to Theorem 8.16.

Theorem 10.39. *Suppose $\{u_n(z)\}$ is a sequence of real-valued harmonic functions that converges uniformly on all compact subsets of a domain D to a function $u(z)$. Then $u(z)$ is harmonic throughout D.*

Proof. Since $u_n(z)$ is continuous for each n, the continuity of $u(z)$ is a consequence of Theorem 6.26. Given $z_0 \in D$ and a disk $|z - z_0| \leq r$ contained in D, we have for each n that

$$u_n(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u_n(z_0 + re^{i\theta})\, d\theta.$$

By Theorem 8.11,

$$\begin{aligned} u(z_0) = \lim_{n\to\infty} u_n(z_0) &= \lim_{n\to\infty} \frac{1}{2\pi} \int_0^{2\pi} u_n(z_0 + re^{i\theta})\, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{i\theta})\, d\theta. \end{aligned}$$

Thus, u has the mean-value property. The result now follows from Theorem 10.38. ∎

Harnack's inequality leads to a theorem concerning sequences of harmonic functions.

Theorem 10.40. (Harnack's Principle) *Suppose $\{u_n(z)\}$ is a sequence of real-valued harmonic functions defined in a domain D, and that $u_{n+1}(z) \geq u_n(z)$ for each $z \in D$ and each n. If $\{u_n(z)\}$ converges for at least one point in D, then $\{u_n(z)\}$ converges for all points in D. Furthermore, the convergence is uniform on compact subsets of D, and the limit function is harmonic throughout D.*

Proof. We may assume that $u_n(z) \geq 0$; for if not, the theorem can be proved for the nonnegative sequence $\{u_n(z) - u_1(z)\}$. By the monotonicity property, for each z in D, either $\{u_n(z)\}$ converges or approaches ∞. Let

$$A = \{z \in D : \, u_n(z) \to \infty\}, \quad \text{and} \quad B = \{z \in D : \, u_n(z) \text{ converges}\}.$$

Given $z_0 \in D$, choose a disk $|z - z_0| \leq R$ contained in D. Then for all z satisfying $|z - z_0| \leq R/2$, Harnack's inequality gives

$$\frac{1}{3} u_n(z_0) = \frac{R - R/2}{R + R/2} u_n(z_0) \leq u_n(z) \leq \frac{R + R/2}{R - R/2} u_n(z_0) = 3u_n(z_0). \quad (10.18)$$

If $u_n(z_0) \to \infty$, the left hand inequality of (10.18) shows that $u_n(z) \to \infty$ for $|z - z_0| \leq R/2$. If $\{u_n(z_0)\}$ converges, the right hand inequality shows that $\{u_n(z)\}$ converges for $|z - z_0| \leq R/2$. Hence, A and B are both open sets, with $A \cup B = D$. Since the domain D is connected, either $A = \emptyset$ or $B = \emptyset$. By hypothesis, there is at least one point in B. Thus $B = D$, and $\{u_n(z)\}$ converges for all z in D.

Next we must show that $\{u_n(z)\}$ converges uniformly on compact subsets of D. Applying Harnack's inequality to $u_{n+p}(z) - u_n(z)$, we get as in (10.18),

$$u_{n+p}(z) - u_n(z) \leq 3[u_{n+p}(z_0) - u_n(z_0)] \quad (10.19)$$

for $|z - z_0| \leq R/2$ and $p = 1, 2, \ldots$. By the Cauchy criterion,

$$u_{n+p}(z_0) - u_n(z_0) < \epsilon \quad (n > N(\epsilon)).$$

Hence from (10.19), we see that $\{u_n(z)\}$ converges uniformly in some neighborhood of z_0. Since z_0 was arbitrary, to every point in D there corresponds a neighborhood in which the convergence of $\{u_n(z)\}$ is uniform.

Now let K be a compact subset of D. For each point of K, construct a neighborhood in which $\{u_n(z)\}$ converges uniformly. By the Heine–Borel theorem, finitely many such neighborhoods cover K. But a sequence converging uniformly on finitely many different sets must converge uniformly on their union. Therefore, $\{u_n(z)\}$ converges uniformly on K.

Finally, it follows from Theorem 10.39 that the limit function is harmonic throughout D. ∎

Remark 10.41. According to Harnack's principle, boundedness of the sequence $\{u_n(z)\}$ at one point forces the boundedness for all other points of D. Further, the contrapositive of the theorem says that if the sequence approaches ∞ at one point in the domain, then it approaches ∞ at all points. That this can actually happen is seen by considering the sequence $u_n(z) = x + n$, which is harmonic in every domain and satisfies the conditions of the theorem. ●

We turn now to the class of analytic functions with positive real part in the disk $|z| < 1$, and apply our knowledge of harmonic functions. According to Harnack's inequality, if $u(z)$ is harmonic and positive for $|z| < 1$ with $u(0) = 1$, then

$$u(z) \leq \frac{1+|z|}{1-|z|} \quad (|z| < 1).$$

Consider the following generalization to analytic functions.

Theorem 10.42. *Suppose $f(z)$ is analytic for $|z| < 1$ with $f(0) = 1$. If $\operatorname{Re} f(z) > 0$ for $|z| < 1$, then*

$$|f(z)| \leq \frac{1+|z|}{1-|z|} \quad (|z| < 1).$$

Proof. Set $\operatorname{Re} f(z) = u(z)$. In view of (10.11), we may write

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{Re^{i\phi}+z}{Re^{i\phi}-z} u(Re^{i\phi})\, d\phi \quad (|z| < R < 1).$$

Hence,

$$|f(z)| \leq \frac{R+|z|}{R-|z|}\frac{1}{2\pi}\int_0^{2\pi} u(Re^{i\phi})\, d\phi = \frac{R+|z|}{R-|z|}u(0) = \frac{R+|z|}{R-|z|}.$$

By letting $R \to 1^-$, the result is obtained.

Here is an alternate proof which relies on Schwarz's inequality (Schwarz's lemma) rather than on Harnack's inequality. If $\operatorname{Re} f(z) > 0$, then the function

$$g(z) = \frac{f(z)-1}{f(z)+1} \tag{10.20}$$

satisfies $|g(z)| < 1$ for $|z| < 1$. Since $g(0) = 0$, it follows from Schwarz's inequality that $|g(z)| \leq |z|$ for $|z| < 1$. Solving for $f(z)$ in (10.20), we get

$$f(z) = \frac{1+g(z)}{1-g(z)}.$$

But

$$|f(z)| \leq \frac{1+|g(z)|}{1-|g(z)|} \leq \frac{1+|z|}{1-|z|},$$

and the proof is complete. ∎

Clearly, Theorem 10.42 is a generalization of Harnack's inequality because $\operatorname{Re} f(z) \leq |f(z)|$.

Remark 10.43. The assumption $f(0) = 1$ does not restrict the generality of the inequality. For if $\operatorname{Re} f(z) > 0$, then the theorem can be applied to the function

$$h(z) = \frac{f(z) - i\operatorname{Im} f(0)}{\operatorname{Re} f(0)},$$

which satisfies the conditions $\operatorname{Re} h(z) > 0, h(0) = 1$. Also, if we had assumed only that $\operatorname{Re} f(z) \geq 0$ for $|z| < 1$, we could have deduced from the open mapping theorem that $\operatorname{Re} f(z) > 0$ for $|z| < 1$. ●

Our next theorem may also be proved by methods that rely on harmonic functions or on Schwarz's lemma.

Theorem 10.44. *Suppose* $f(z) = 1 + \sum_{n=1}^{\infty} a_n z^n$ *is analytic for* $|z| < 1$*. If* $\operatorname{Re} f(z) > 0$ *for* $|z| < 1$*, then* $|a_n| \leq 2$ *for every* n*.*

Proof. Set $f(z) = u(re^{i\theta}) + iv(re^{i\theta})$, with $a_n = \alpha_n + i\beta_n$. Then

$$u(re^{i\theta}) = 1 + \operatorname{Re} \sum_{m=1}^{\infty} a_m r^m e^{im\theta} = 1 + \sum_{m=1}^{\infty} (\alpha_m \cos m\theta - \beta_m \sin m\theta) r^m.$$

This series converges uniformly on the circle $|z| = r < 1$. By Theorem 8.11, we may multiply by $\cos n\theta$ or $\sin n\theta$ and then integrate term-by-term. Since

$$\int_0^{2\pi} \cos n\theta \cos m\theta \, d\theta = \int_0^{2\pi} \sin n\theta \sin m\theta \, d\theta = 0$$

for $n \neq m$ and

$$\int_0^{2\pi} \cos n\theta \sin m\theta \, d\theta = 0$$

for all n and m, we have the identities

$$\frac{1}{\pi} \int_0^{2\pi} u(re^{i\theta}) \cos n\theta \, d\theta = \frac{1}{\pi} \int_0^{2\pi} \alpha_n r^n \cos^2 n\theta \, d\theta = \alpha_n r^n, \tag{10.21}$$

$$\frac{1}{\pi} \int_0^{2\pi} u(re^{i\theta}) \sin n\theta \, d\theta = \frac{1}{\pi} \int_0^{2\pi} -\beta_n r^n \sin^2 n\theta \, d\theta = -\beta_n r^n. \tag{10.22}$$

Multiplying (10.22) by $-i$ and adding to (10.21), we obtain

$$a_n r^n = (\alpha_n + i\beta_n) r^n = \frac{1}{\pi} \int_0^{2\pi} u(re^{i\theta}) e^{-in\theta} \, d\theta.$$

Thus,

$$|a_n|r^n \le \frac{1}{\pi}\int_0^{2\pi} u(re^{i\theta})|e^{-in\theta}|\,d\theta = \frac{1}{\pi}\int_0^{2\pi} u(re^{i\theta})\,d\theta.$$

By the mean-value property,

$$\frac{1}{\pi}\int_0^{2\pi} u(re^{i\theta})\,d\theta = 2u(0) = 2.$$

Hence, $|a_n|r^n \le 2$. Letting $r \to 1^-$, the result follows.

As an alternate proof, consider

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

which is analytic in the disk $|z|<1$, with $g(0)=0$ and $|g(z)|<1$. By Exercise 8.73(10), $|g'(0)| \le 1$. But $g'(0) = a_1/2$, so that $|a_1| \le 2$.

We will now show that $|a_n| \le 2$, for arbitrary n by constructing a new function of the form $1 + a_n z + \cdots$, which satisfies the conditions of the theorem. In view of the identity

$$\sum_{k=1}^{n} e^{(2k\pi i)m/n} = \begin{cases} n & \text{if } m \text{ is a multiple of } n \\ 0 & \text{otherwise,} \end{cases}$$

we can verify (do it!) that the function

$$h(z) = \frac{1}{n}\sum_{k=1}^{n} f(e^{2k\pi i/n} z^{1/n}) = 1 + a_n z + \cdots$$

is analytic for $|z|<1$, and has positive real part. Therefore, $|a_n| \le 2$ and the proof is complete. ■

The function

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

maps the circle $|z|=1$ onto the imaginary axis and the disk $|z|<1$ onto the right half-plane. This function shows that equality holds in the previous two theorems. That is,

$$\operatorname{Re} f(z) = \frac{1+|z|}{1-|z|}$$

when z is a positive real number, and

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

Questions 10.45.

1. Can the mean-value property hold for discontinuous functions?

2. If two continuous functions satisfy the mean-value property, does their sum? Their product?
3. Suppose $\{u_n(z)\}$ is a sequence of harmonic functions having harmonic conjugates $\{v_n(z)\}$. If $\{u_n(z)\}$ converges uniformly in a region, does $\{v_n(z)\}$ also converge uniformly?
4. Is the conclusion of Theorem 10.40 valid if the hypothesis $u_{n+1}(z) \geq u_n(z)$ is replaced with $u_{n+1}(z) \leq u_n(z)$?
5. What kind of generalizations of Theorem 10.42 can you prove?
6. Why are some theorems valid for compact subsets of a domain but not for the whole domain?
7. Where was the positivity of $\operatorname{Re} f(z)$ used in the first proof of Theorem 10.44?
8. What is the relationship between Schwarz's inequality and Harnack's inequality?

Exercises 10.46.

1. Suppose $\{u_n(z)\}$ is a sequence of functions harmonic in a domain D, and that $u_{n+1}(z) \geq u_n(z)$ for each $z \in D$ and each n. If $u_n(z_0) \to \infty$ for some $z_0 \in D$, show that $u_n(z) \to \infty$ uniformly on compact subsets. That is, given a compact subset C and a real number M, show that $u_n(z) \geq M$ for $n > N$ and all $z \in C$.
2. Let K be a compact subset of a domain D. Given $z_0 \in D$, show that there exist real constants A and B (depending on z_0, K, and D) such that
$$A \cdot u(z_0) \leq u(z) \leq B \cdot u(z_0)$$
for all z in K and all functions $u(z)$ harmonic in D.
3. A continuous real-valued function $u(z)$ is said to be subharmonic in a domain in D if
$$u(z_0) \leq \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{i\theta})\, d\theta$$
for every disk $|z - z_0| \leq r$ contained in D. Show that a nonconstant subharmonic function cannot attain a maximum in a domain. Can it attain a minimum?
4. Suppose $f(z)$ is analytic with $\operatorname{Re} f(z) > 0$ for $|z| < 1$. If $f(0) = 1$, then apply Theorem 10.42 to $1/f(z)$ to show that
$$|f(z)| \geq \frac{1 - |z|}{1 + |z|}.$$
Can this be deduced from Harnack's inequality?
5. Suppose $g(z)$ is analytic for $|z| < 1$ with $g(0) = 1$. If $\operatorname{Re} g(z) > \alpha$, show that
$$|g(z)| \leq \frac{1 + (1 - 2\alpha)|z|}{1 - |z|} \quad (|z| < 1).$$

6. Under the assumptions of the previous exercise, show that

$$|g(z)| \geq \frac{1-(1-2\alpha)|z|}{1+|z|} \quad (|z|<1).$$

7. Suppose that $g(z) = 1 + \sum_{n=1}^{\infty} a_n z^n$ is analytic for $|z| < 1$, with $\operatorname{Re} g(z) > \alpha$. Show that $|a_n| \leq 2(1-\alpha)$ for every n.
8. Suppose that $g(z)$ is analytic for $|z| < 1$ with $g(0) = a > 0$. If $\operatorname{Re} g(z) > 0$ in $|z| < 1$, then show that

$$\left|\frac{g(z)-a}{g(z)+a}\right| \leq |z| \text{ for } |z|<1, \quad \text{and} \quad |g'(0)| \leq 2a.$$

9. Suppose that $g(z)$ is analytic for $|z| < 1$, and $g(0) = 0$. If $\operatorname{Re} g(z) < a$ $(a > 0)$, show that

$$|g(z)| \leq \frac{2a|z|}{1-|z|} \text{ for } |z|<1.$$

   Is $|g'(0)| \leq 2a$?
10. Show that equality holds in Theorem 10.42 and Theorem 10.44 if and only if $f(z)$ is of the form

$$f(z) = \frac{1+e^{i\theta_0}z}{1-e^{i\theta_0}z} \quad (\theta_0 \text{ real}).$$

11

Conformal Mapping and the Riemann Mapping Theorem

Our study of mapping properties in Chapters 3 and 4 was limited because derivatives had not yet been introduced. That remedied, we look anew at some old functions. We shall see that the derivative relates the angle between two curves to the angle between their images. In addition, the derivative will be seen to measure the "distortion" of image curves.

Analytic functions mapping disks and half-planes onto disks and half-planes, disks onto the interior of ellipses, etc., have previously been constructed. The major result of this chapter, known as the Riemann mapping theorem, tells us that there is nearly always an analytic function that maps a given simply connected domain onto another given simply connected domain. This is a very powerful result and is used in a wide range of mathematical settings. Our method of proof relies on normal families, a concept that enables us to extract limit functions from families of functions. Recall how we previously had extracted limit points from sequences of points (Bolzano–Weierstrass theorem).

11.1 Conformal Mappings

Any straight line in the plane that passes through the origin may be parameterized by $\sigma(s) = se^{i\alpha}$, where s traverses the set of real numbers and α is the angle−measured in radians−between the positive real axis and the line. More generally, a straight line passing through the point z_0 and making an angle α with the real axis can be expressed as $\sigma(s) = z_0 + se^{i\alpha}$, s real.

Suppose now that a function f is analytic on a smooth (parameterized) curve $z(t)$, $t \in [a, b]$. Then the image of $z(t)$ under f is also a smooth curve whose derivative is given by $f'(z(t))z'(t)$. A smooth curve is characterized by having a tangent at each point. So, we interpret $z'(t)$ as a vector in the direction of the tangent vector at the point $z(t)$. Our purpose is to compare the inclination of the tangent to the curve at a point with the inclination of the tangent to the image curve at the image of the point.

Let $z_0 = z(t_0)$ be a point on the curve $z = z(t)$. Then the vector $z'(t_0)$ is tangent to the curve at the point z_0 and $\arg z'(t_0)$ is the angle this directed tangent makes with the positive x-axis. Suppose that $w = w(t) = f(z(t))$, with $w_0 = f(z_0)$. For any point z on the curve other than z_0, we have the identity

$$w - w_0 = \frac{f(z) - f(z_0)}{z - z_0}(z - z_0).$$

Thus,

$$\arg(w - w_0) = \arg \frac{f(z) - f(z_0)}{z - z_0} + \arg(z - z_0) \pmod{2\pi}, \tag{11.1}$$

where it is assumed that $f(z) \neq f(z_0)$ so that (11.1) has meaning. Note that $\arg(z - z_0)$ is the angle in the z plane between the x axis and the straight line passing through the points z and z_0, while $\arg(w - w_0)$ is the angle in the w plane between the u axis and the straight line passing through the points w and w_0. Hence as z approaches z_0 along the curve $z(t)$, $\arg(z - z_0)$ approaches a value θ, which is the angle that the tangent to the curve $z(t)$ at z_0 makes with the x axis. Similarly, $\arg(w - w_0)$ approaches a value ϕ, the angle that the tangent to the curve $f(z(t))$ at w_0 makes with the u axis.

Suppose $f'(z_0) \neq 0$ so that $\arg f'(z_0)$ has meaning. Then taking limits in (11.1), we find (mod 2π) that

$$\phi = \arg f'(z_0) + \theta, \quad \text{or} \quad \arg w'(t_0) = \arg f'(z_0) + \arg z'(t_0). \tag{11.2}$$

That is, the difference between the tangent to a curve at a point and the tangent to the image curve at the image of the point depends only on the derivative of the function at the point (see Figure 11.1).

For instance, consider $f(z) = z^2$. Then $f'(z) \neq 0$ on $\mathbb{C}\setminus\{0\}$. Choose $z_0 = 1 + i$. Then $f'(z_0) = 2(1 + i)$ so that

$$\arg f'(z_0) = (\pi/4) + 2k\pi.$$

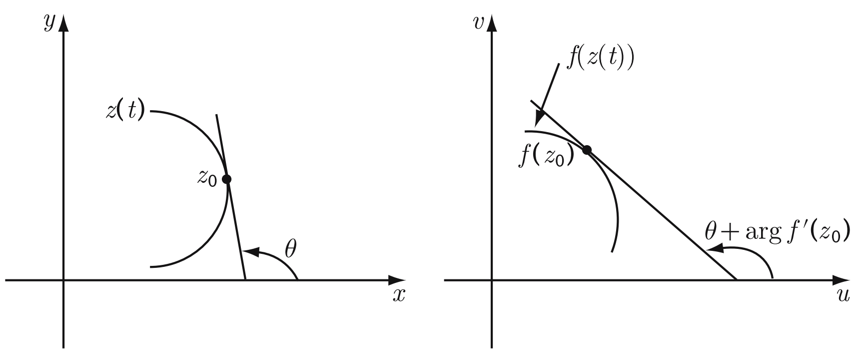


Figure 11.1. The direction of the tangent line at $z(t)$

To verify the angle of rotation of a particular curve, we consider a simple curve C passing through z_0:

$$C: z(t) = t(1+i), \quad t \in \mathbb{R}.$$

Clearly, $\pi/4$ is the angle which the curve C makes with the x axis. The image of C under $f(z) = z^2 = (x^2 - y^2) + i(2xy)$ is given by $w(t) = 0 + 2t^2 i$. Thus, the angle of rotation at $1+i$ is $\pi/2$ which corresponds to the case $k = 0$.

If two smooth curves intersect at a point, then the angle between these two curves is defined as the angle between the tangents to these curves at the point. We can now state

Theorem 11.1. *Suppose $f(z)$ is analytic at z_0 with $f'(z_0) \neq 0$. Let $C_1 : z_1(t)$ and $C_2 : z_2(t)$ be smooth curves in the z plane that intersect at $z_0 =: z_1(t_0) =: z_2(t_0)$, with $C_1' : w_1(t)$ and $C_2' : w_2(t)$ the images of C_1 and C_2, respectively. Then the angle between C_1 and C_2 measured from C_1 to C_2 is equal to the angle between C_1' and C_2' measured from C_1' to C_2'.*

Proof. Let the tangents to C_1 and C_2 make angles θ_1 and θ_2, respectively, with the x axis (see Figure 11.2). Then the angle between C_1 and C_2 is $\theta_2 - \theta_1$.

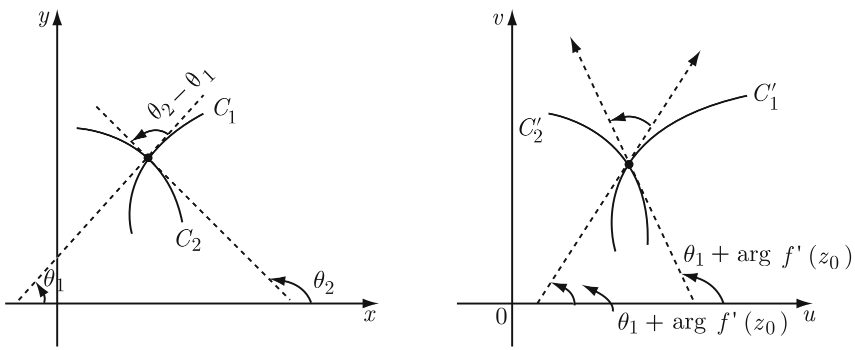


Figure 11.2. The curves C_1 and C_2 intersect at angle α

According to (11.2), the angle between C_1' and C_2', which is the angle between the tangent vectors $f'(z_0)z_1'(t_0)$ and $f'(z_0)z_2'(t_0)$, of the image curves is

$$\theta_2 + \arg f'(z_0) - (\theta_1 + \arg f'(z_0)\ ) = \theta_2 - \theta_1,$$

and the theorem is proved. ■

A function that preserves both angle size and orientation is said to be *conformal*. Theorem 11.1 says that an analytic function is conformal at all points where the derivative is nonzero. We have already discussed a number of examples of conformal maps without referring to the name "conformal".

For instance, $f(z) = e^z$ maps vertical and horizontal lines into circles and orthogonal radial rays, respectively.

A function that preserves angle size but not orientation is said to be *isogonal*. An example of such a function is $f(z) = \overline{z}$. To illustrate, $\overline{z}$ maps the positive real axis and the positive imaginary axis onto the positive real axis and the negative real axis respectively (see Figure 11.3). Although the two curves intersect at right angles in each plane, a "counterclockwise" angle is mapped onto a "clockwise" angle.

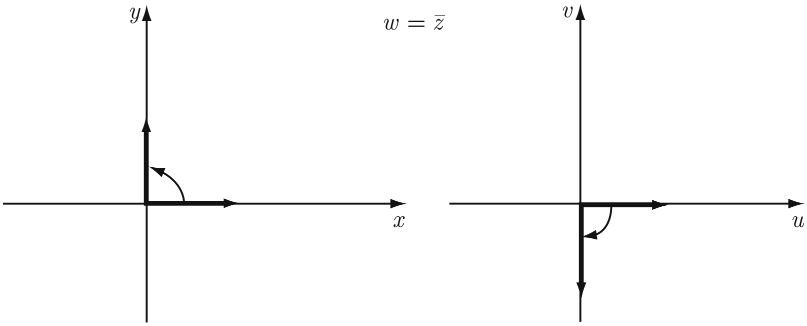


Figure 11.3.

Suppose $f(z)$ is analytic at z_0 and $f'(z_0) \neq 0$. When z is near z_0, there is an interesting relationship concerning the distance between the points z and z_0 and the distance between their images. Note that

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \epsilon(z)(z - z_0)$$

where $\epsilon(z) \to 0$ as $z \to z_0$. Thus for z close to z_0,

$$f(z) \approx f'(z_0)z + (-f'(z_0)z_0 + f(z_0))$$

so that we may approximate $f(z)$ by the linear function. Also,

$$|f(z) - f(z_0)| \approx |f'(z_0)|\,|z - z_0|\,. \tag{11.3}$$

In view of (11.3), "small" neighborhoods of z_0 are mapped onto roughly the same configuration, magnified by the factor $|f'(z_0)|$, see Figure 11.4. Hence, $f'(z_0)$ plays two roles in determining the geometric character of the image. According to (11.2), $\arg f'(z_0)$ measures the rotation; according to (11.3), $|f'(z_0)|$ measures (for points nearby) the magnification or distortion of the image.

An interesting comparison can now be made between the derivatives of real and complex functions. For real differentiable functions, the nonvanishing of the derivative is sufficient to guarantee that the function is one-to-one on an interval. This is not the case for complex functions on a domain. Even though

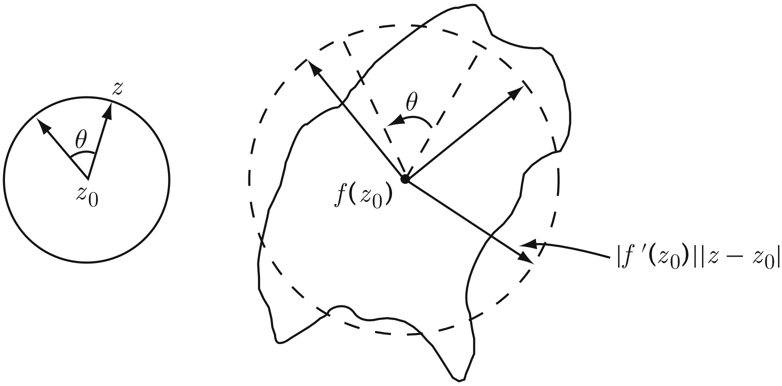


Figure 11.4.

the derivative of the entire function e^z never vanishes, we have $e^z = e^{z+2\pi i}$ for all z. Similarly, the entire function $f(z) = z^2$ is conformal on $\mathbb{C}\setminus\{0\}$. However, it is geometrically intuitive (Figure 11.4) that the nonvanishing of a derivative implies, at least locally, that the function is one-to-one. We now show this formally in the following form which gives a sufficient condition for the existence of a local inverse.

Theorem 11.2. *If $f(z)$ is analytic at z_0 with $f'(z_0) \neq 0$, then $f(z)$ is one-to-one in some neighborhood of z_0.*

Proof. Since $f'(z_0) \neq 0$ and $f'(z)$ is continuous at z_0, there exists a $\delta > 0$ such that

$$|f'(z) - f'(z_0)| < \frac{|f'(z_0)|}{2} \quad \text{for all } |z| < \delta.$$

Let z_1 and z_2 be two distinct points in $|z| < \delta$, and γ be a line segment connecting z_1 and z_2. Set $\phi(z) = f(z) - f'(z_0)z$ so that $|\phi'(z)| < |f'(z_0)|/2$ for all $|z| < \delta$. Now we have

$$|\phi(z_2) - \phi(z_1)| = \left| \int_\gamma \phi'(z)\, dz \right| < (|f'(z_0)|/2)|z_2 - z_1|,$$

or equivalently,

$$|f(z_2) - f(z_1) - f'(z_0)(z_2 - z_1)| < (|f'(z_0)|/2)|z_2 - z_1|.$$

Thus, by the triangle inequality, we obtain

$$|f(z_2) - f(z_1)| > (|f'(z_0)|/2)|z_2 - z_1| > 0.$$

That is, $f(z)$ is one-to-one in $|z| < \delta$. ∎

The vanishing of a derivative does not preclude the possibility of a real function being one-to-one. Although the derivative of $f(x) = x^3$ is zero at the origin, the function is still one-to-one on the real line. That this cannot occur for complex functions is seen by

Theorem 11.3. *If $f(z)$ is analytic and one-to-one in a domain D, then $f'(z) \neq 0$ in D, so that f is conformal on D.*

Proof. If $f'(z) = 0$ at some point z_0 in D, then

$$f(z) - f(z_0) = \frac{f''(z_0)}{2!}(z - z_0)^2 + \cdots$$

has a zero of order k $(k \geq 2)$ at z_0. Since zeros of an analytic function are isolated, there exists an $r > 0$ so small that both $f(z) - f(z_0)$ and $f'(z)$ have no zeros in the punctured disk $0 < |z - z_0| \leq r$. Let $g(z) := f(z) - f(z_0)$, $C = \{z : |z - z_0| = r\}$ and

$$m = \min_{z \in C} |g(z)|.$$

Then, g has a zero of order k $(k \geq 2)$ and $m > 0$. Let $b \in \mathbb{C}$ be such that $0 < |b - f(z_0)| < m$. Then, as $m \leq |g(z)|$ on C,

$$|f(z_0) - b| < |g(z)| \quad \text{on } C.$$

It follows from Rouche's theorem that $g(z)$ and

$$g(z) + (f(z_0) - b) = f(z) - b$$

have the same number of zeros inside C. Thus, $f(z) - b$ has at least two zeros inside C. Observe that none of these zeros can be at z_0. Since $f'(z) \neq 0$ in the punctured disk $0 < |z - z_0| \leq r$, these zeros must be simple and so, distinct. Thus, $f(z) = b$ at two or more points inside C. This contradicts the fact that f is one-to-one on D. ■

We sum up our results for differentiable functions. In the real case, the nonvanishing of a derivative on an interval is a sufficient but not a necessary condition for the function to be one-to-one on the interval; whereas in the complex case, the nonvanishing of a derivative on a domain is a necessary but not a sufficient condition for the function to be one-to-one on the domain.

An analytic function $f : D \to \mathbb{C}$ is called *locally bianalytic* at $z_0 \in D$ if there exists a neighborhood N of z_0 such that restriction of f from N onto $f(N)$ is bianalytic. Clearly, a locally bianalytic map on D need not be bianalytic on D, as the example $f(z) = z^n$ $(n > 2)$ on $\mathbb{C}\backslash\{0\}$ illustrates.

Combining Theorem 11.2 and Theorem 11.3 leads to the following criterion for local bianalytic maps.

Theorem 11.4. *Let $f(z)$ be analytic in a domain D and $z_0 \in D$. Then f is bianalytic at z_0 iff $f'(z_0) \neq 0$.*

A sufficient condition for an analytic function to be one-to-one in a simply connected domain is that it be one-to-one on its boundary. More formally, we have

Theorem 11.5. *Let $f(z)$ be analytic in a simply connected domain D and on its boundary, the simple closed contour C. If $f(z)$ is one-to-one on C, then $f(z)$ is one-to-one in D.*

Proof. Choose a point z_0 in D such that $w_0 = f(z_0) \neq f(z)$ for z on C. According to the argument principle, the number of zeros of $f(z) - f(z_0)$ in D is given by $(1/2\pi)\triangle_C \{f(z) - f(z_0)\}$. By hypothesis, the image of C must be a simple closed contour, which we shall denote by C' (see Figure 11.5). Thus the net change in the argument of $w - w_0 = f(z) - f(z_0)$ as $w = f(z)$ traverses the contour C' is either $+2\pi$ or -2π, according to whether the contour is traversed counterclockwise or clockwise. Since $f(z)$ assumes the value w_0 at least once in D, we must have

$$\frac{1}{2\pi}\triangle_C \{f(z) - f(z_0)\} = \frac{1}{2\pi}\triangle_C \{w - w_0\} = 1.$$

That is, $f(z)$ assumes the value $f(z_0)$ exactly once in D.

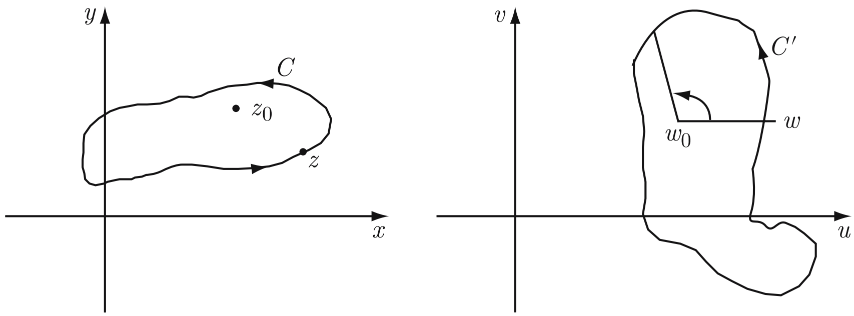


Figure 11.5.

This proves the theorem for all points z_0 in D at which $f(z) \neq f(z_0)$ when z is on C. If $f(z) = f(z_0)$ at some point on C, then the expression $\triangle_C \{f(z) - f(z_0)\}$ is not defined. We leave for the reader the completion of the proof in this special case. ■

In the proof of Theorem 11.1, we relied on the nonvanishing of the derivative. In Theorem 11.2, we see that every analytic function is locally one-to-one at points where the derivative is nonvanishing. More generally, it can be shown that if f is analytic at z_0 and f' has a zero of order k at z_0, then f is locally $(k+1)$-to-one. For example, if $f(z) = z^2$, then $f'(z)$ has a zero of order 1 at the origin and hence, it is two-to-one in any neighborhood of the origin.

We now examine the behavior of an analytic function in a neighborhood of a critical point, a point where the derivative vanishes. First we note that the angle of intersection of two smooth curves at a critical point of an analytic function is not the same as the angle of intersection of their images under f. If $f(z)$ is analytic and $f'(z)$ has a zero of order $k-1$ at $z = z_0$, then $f^{(j)}(z) = 0$ for $j = 1, \cdots, k-1$ and so we may write

$$f(z) = f(z_0) + a_k(z - z_0)^k + a_{k+1}(z - z_0)^{k+1} + \cdots .$$

Thus, $f(z) - f(z_0) = (z - z_0)^k g(z)$, where $g(z)$ is analytic at z_0 and $g(z_0) = a_k \neq 0$. Consequently,

$$\arg[f(z) - f(z_0)] = k \arg(z - z_0) + \arg g(z). \tag{11.4}$$

Suppose θ is the angle that the tangent to a smooth curve C at z_0 makes with the x axis, and ϕ is the angle that the tangent to the image C' of the curve C at $f(z_0)$ makes with the u axis. If z approaches z_0 along the curve C, then $w = f(z)$ approaches $w_0 = f(z_0)$ along the curve C', and so (11.4) yields

$$\phi = k\theta + \arg g(z_0). \tag{11.5}$$

Observe that (11.5) reduces to (11.2) in the special case when $k = 1$. In general, the tangent to an image curve depends on the tangent to the original curve as well as on the order and argument of the first nonzero derivative at the point in question. Just as (11.2) led to Theorem 11.1, so (11.5) leads to

Theorem 11.6. *Suppose $f(z)$ is analytic at z_0, and that $f'(z)$ has a zero of order $k-1$ at z_0. If two smooth curves in the domain of f intersect at an angle θ, then their images intersect at an angle $k\theta$.*

Proof. Suppose that the tangents to the two curves make angles θ_1 and θ_2 with respect to the real axis. Then $\theta = \theta_2 - \theta_1$ is the angle between the two curves. According to (11.5), the angle ϕ between their images is given by

$$\phi = k\theta_2 + \arg g(z_0) - (k\theta_1 + \arg g(z_0)) = k\theta, \quad g(z_0) = \frac{f^{(k)}(z_0)}{k!}. \qquad \blacksquare$$

Combining Theorems 11.1 and 11.6, we see that an analytic function is conformal at a point if and only if it has a nonzero derivative at the point. Thus, an analytic function f is conformal on a domain D iff $f'(z) \neq 0$ on D.

It now pays to reexamine bilinear transformations, studied in Chapter 3, from a conformal mapping point of view. Recall that the transformation

$$w = f(z) = \frac{az + b}{cz + d} \quad (ad - bc \neq 0) \tag{11.6}$$

represents a one-to-one continuous mapping from the extended plane onto itself, with $f(-d/c) = \infty$ and $f(\infty) = a/c$. Since $f'(z) \neq 0$ $(ad - bc \neq 0)$, the mapping is conformal for all finite z, $z \neq -d/c$.

As we have seen, a circle or a straight line is mapped onto either a circle or a straight line, depending on which point is mapped onto the point at ∞. For instance, the inversion transformation $w = 1/z$ maps straight lines not passing through the origin onto circles. In particular, the lines $y = x + 1$ and $y = -x + 1$ are mapped, respectively, onto the circles

$$\left(u+\frac{1}{2}\right)^2+\left(v+\frac{1}{2}\right)^2=\left(\frac{1}{\sqrt{2}}\right)^2 \quad \text{and} \quad \left(u-\frac{1}{2}\right)^2+\left(v+\frac{1}{2}\right)^2=\left(\frac{1}{\sqrt{2}}\right)^2.$$

At first glance, Figure 11.6 is somewhat misleading. It shows a pair of straight lines that intersect at one point being mapped onto a pair of circles that intersect at two points. It should not be forgotten, however, that these straight lines also intersect at ∞. For both lines, the point $(0, 1)$ is mapped onto the point $(0, -1)$ while the point at ∞ is mapped onto the origin. The two lines intersect at right angles at $(0, 1)$ as do the two circles at $(0, -1)$. This is in harmony with Theorem 11.1.

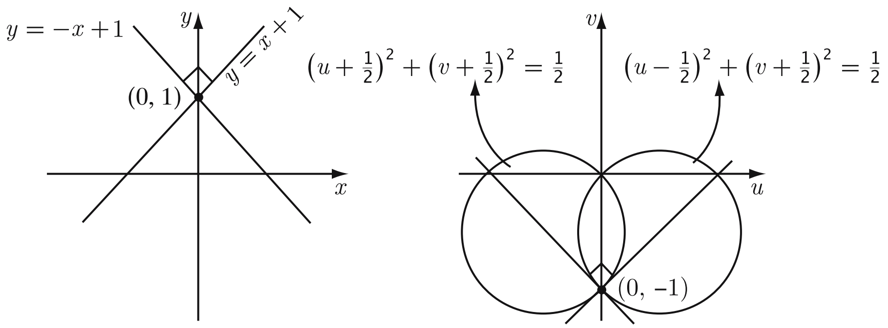


Figure 11.6.

But at what angle do the two lines intersect at ∞? We need the following definition: Two smooth curves in the extended plane are said to intersect at an angle α at ∞ if their images under the transformation $w = 1/z$ intersect at an angle α at the origin. Since the two circles in Figure 11.6 intersect at right angles at the origin, the lines $y = x+1$ and $y = -x+1$ intersect at right angles at ∞.

With this definition, we can show that all transformations of the form (11.6) are conformal at ∞. There are two cases to consider.

Case 1: Let $c \neq 0$. The behavior of f at ∞ is determined from the behavior of $f(1/z)$ at 0 in (11.6). Thus we consider

$$g(z) = f\left(\frac{1}{z}\right) = \frac{a/z + b}{c/z + d} = \frac{bz + a}{dz + c}.$$

Since $g'(0) = (bc - ad)/c^2 \neq 0$, it follows that $g(z)$ is conformal at $\zeta = 0$. But this means that $f(z)$ is conformal at $z = \infty$.

Case 2: Let $c = 0$. Then (11.6) is linear, and maps $z = \infty$ onto $w = \infty$. So we need to consider the expression $h(z) = 1/f(1/z)$ in (11.6):

$$w = h(z) = \frac{dz}{bz + a}.$$

Since $h'(0) = d/a \neq 0$, $h(z)$ is conformal at $z = 0$; that is, $f(z)$ is conformal at $z = \infty$. Hence, *a bilinear transformation is a one-to-one conformal mapping of the extended plane onto itself.*

Recall from Chapter 4 that the exponential function e^z maps lines parallel to the y axis onto circles centered at the origin and lines parallel to the x axis onto rays emanating from the origin. From elementary geometry we know that these two image curves must intersect at right angles (see Figure 11.7).

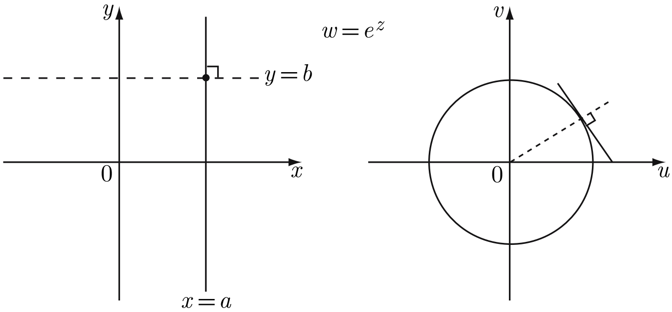


Figure 11.7.

Finally, consider the function $w = \cos z$, which maps lines parallel to the y axis onto ellipses and lines parallel to the x axis onto hyperbolas. According to Theorem 11.1, these conic sections must intersect at right angles (see Figure 11.8).

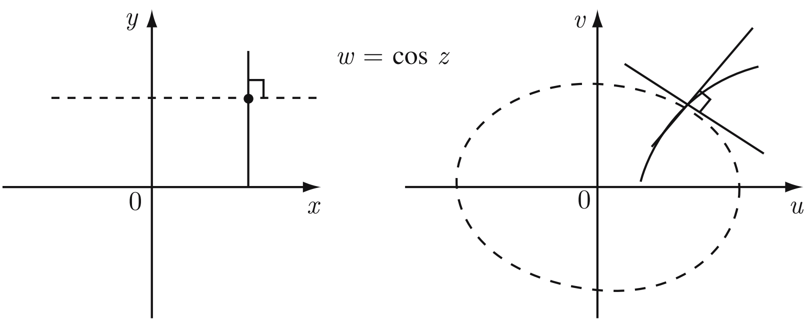


Figure 11.8.

Questions 11.7.

1. What is meant by a tangent to a point on a straight line?
2. Was it necessary to require the curves in Theorem 11.1 to be smooth?
3. Can nonanalytic functions be conformal?
4. What kind of functions are isogonal?
5. Why does the derivative play such a central role?
6. If a function is one-to-one in some neighborhood of each point in a domain, why does this not mean that the function is one-to-one in the domain?
7. If f is conformal on a domain D, is f always one-to-one on D?
8. If f is conformal on a domain D which is symmetric with respect to the real axis, is $\overline{f(\overline{z})}$ conformal on D?
9. What is the relationship between conformal and one-to-one?
10. At what angle do parallel lines intersect at ∞?
11. How might we define a function to be analytic at ∞?
12. Is the sum of conformal maps conformal? The product? The composition?

Exercises 11.8.

1. Given a complex number z_0 and an $\epsilon > 0$, show that there exists a function $f(z)$ analytic at z_0 with $f'(z_0) \neq 0$ and such that $f(z)$ is not one-to-one for $|z - z_0| < \epsilon$. Does this contradict Theorem 11.2?
2. Show that z^2 is one-to-one in a domain D if and only if D is contained in a half-plane whose boundary passes through the origin.
3. Find points at which the mapping defined by $f(z) = nz + z^n$ $(n \in \mathbb{N})$ is not conformal.
4. Prove that two smooth curves intersect at an angle α at ∞ if and only if their images under stereographic projection (see Section 2.4) intersect at an angle α at the north pole.
5. Show that $f(\overline{z})$ and $\overline{f(z)}$ are both isogonal at points where $f(z)$ is analytic with nonzero derivative.
6. If two straight lines are mapped by a bilinear transformation onto circles tangent to each other, show that the two lines must be parallel. Is the converse true?
7. Find the radius of the largest disk centered at the origin in which $w = e^z$ is one-to-one. Is the radius different if the disk is centered at an arbitrary point z_0?
8. For $f(z) = e^z$, find $\arg f'(z)$. Use this to verify that lines parallel to the y axis and x axis map, respectively, onto circles and rays.
9. Suppose $f(z)$ is analytic at z_0 with $f'(z_0) \neq 0$. Prove that a "small" rectangle containing z_0 and having area A is mapped onto a figure whose area is approximately $|f'(z_0)|^2 A$.
10. Either directly or by making use of Theorem 11.5, show that the function $w = z^n$ maps the ray $\arg z = \theta$ $(0 \le \theta < 2\pi/n)$ onto the ray $\arg z = n\theta$.

11. If $f(z)$ is nonconstant and analytic in a domain D, show that $f'(z) = 0$ for only a countable number of points in D. Thus conclude that $f(z)$ is locally one-to-one and conformal at all but a countable number of points in D.
12. Show that $f(z) = z + 1/z$ is conformal except at $z = \pm 1$. With this in mind, review its mapping properties from Chapter 3.

11.2 Normal Families

We have previously seen significant differences between pointwise and uniform continuity as well as between pointwise and uniform convergence. Once again we encounter the contrast between local and global properties. This time, we shall require a uniformity to hold over a set consisting of a family of functions.

A family $\mathcal{F}$ of functions is said to be *uniformly bounded* on a set A if there exists a real number M such that $|f(z)| \leq M$ for all $f \in \mathcal{F}$ and all $z \in A$. Certainly the uniform boundedness of a family implies that each member of the family is bounded. On the other hand, each member of the sequence $\{f_n(z)\}$ of functions $f_n(z) = nz$ is bounded in the disk $|z| \leq R$, but there is no bound that works for every member of the family.

A family $\mathcal{F}$ of functions is said to be *locally uniformly bounded* on a set A if to each $z \in A$ there corresponds a neighborhood in which $\mathcal{F}$ is uniformly bounded. The sequence $f_n(z) = 1/(1 - z^n)$ is locally uniformly bounded, but not uniformly bounded in the disk $|z| < 1$. We have the following characterization:

Theorem 11.9. *A family $\mathcal{F}$ of functions is locally uniformly bounded in a domain D if and only if $\mathcal{F}$ is uniformly bounded on each compact subset of D.*

Proof. Let $\mathcal{F}$ be locally uniformly bounded and suppose K is a compact subset of D. For each point in K, choose a neighborhood in which $\mathcal{F}$ is uniformly bounded. This provides an open cover for K. According to the Heine–Borel theorem, there exists a finite subcover of K. That is, there are finitely many $z_i \in K$ and $\epsilon_i > 0$ such that $K \subset \bigcup_{i=1}^{n} N(z_i; \epsilon_i)$, where $|f(z)| \leq M_i$ for all $f \in \mathcal{F}$ and all $z \in N(z_i; \epsilon_i)$. Then $\mathcal{F}$ is uniformly bounded on K, having for a bound $M = \max\{M_1, M_2, \ldots, M_n\}$.

The converse is immediate from the fact that the closure of a neighborhood of a point is a compact set. ∎

By restricting ourselves to locally uniformly bounded families of analytic functions, we can obtain additional information.

Theorem 11.10. *Suppose $\mathcal{F}$ is a family of locally uniformly bounded analytic functions in a domain D. Then the family $\mathcal{F}^{(n)}$, consisting of the nth derivatives of all functions in $\mathcal{F}$, is also locally uniformly bounded in D.*

Proof. It suffices to prove this when $n = 1$, since then the result may be reapplied successively to each new class. Suppose for some z_0 in D that $|f(z)| \le M$ for each $f \in \mathcal{F}$ and all z inside or on the circle $C : |z - z_0| = r$ contained in D. Then for z in the smaller disk $|z - z_0| \le r/2$, Cauchy's integral formula yields

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

and so, as $|\zeta - z| \ge |\zeta - z_0| - |z - z_0| \ge r - r/2 = r/2$,

$$|f'(z)| \le \frac{1}{2\pi (r/2)^2} \int_C |f(\zeta)|\, |d\zeta| \le \frac{4M}{r}.$$

This shows that the family $\mathcal{F}'$ is locally uniformly bounded at z_0. Since z_0 was arbitrary, the proof is complete. ∎

We next extend the concept of uniform continuity. A family $\mathcal{F}$ of functions is said to be *equicontinuous* in a region R if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(z_1) - f(z_0)| < \epsilon$ for all $f \in \mathcal{F}$ and all points $z_0, z_1 \in R$ satisfying $|z_1 - z_0| < \delta$. Observe that each member of an equicontinuous family is uniformly continuous. That is, for an equicontinuous family we can find a $\delta = \delta(\epsilon)$ that works for all points in the set as well as for all functions in the family.

It is possible for each member of a family to be uniformly continuous without the family being equicontinuous. To see this, set $f_n(z) = nz$. Each f_n is uniformly continuous on $|z| \le R$ because

$$|f_n(z_1) - f_n(z_0)| = n|z_1 - z_0| < \epsilon$$

whenever $|z_1 - z_0| < \epsilon/n = \delta$. But a δ cannot be chosen that works for all n. Hence the sequence $\{nz\}$ is not equicontinuous on $|z| \le R$.

There is an important relationship between locally uniformly bounded and equicontinuous families of analytic functions.

Theorem 11.11. *If $\mathcal{F}$ is a locally uniformly bounded family of analytic functions in a domain D, then $\mathcal{F}$ is equicontinuous on compact subsets of D.*

Proof. We prove the theorem in the special case that K is a closed disk contained in D. The proof for general compact subsets of D is similar to the proof of Theorem 11.9, and is left for the reader. By Theorem 11.10, the family $\mathcal{F}'$, consisting of the derivatives of functions in $\mathcal{F}$, is also locally uniformly bounded. In view of Theorem 11.9. we may therefore assume that $|f'(z)| \le M$ for all $f \in \mathcal{F}$ and all $z \in K$. Then for $z_0, z_1 \in K$, we have

$$|f(z_1) - f(z_0)| = \left| \int_{z_0}^{z_1} f'(z)\, dz \right| \le M|z_1 - z_0|,$$

where the path from z_0 to z_1 is taken to be the straight line segment. By choosing $\delta = \epsilon/M$ (ϵ arbitrary), we see that the family $\mathcal{F}$ is equicontinuous on the disk K. ∎

Remark 11.12. The converse of Theorem 11.11 is not true. The sequence $f_n(z) = z + n$ is equicontinuous on all compact subsets of the plane. In fact $f_n(z_1) - f_n(z_0) = z_1 - z_0$ for each n, so that $\delta = \epsilon$ may be chosen. However, $\{f_n(z)\}$ is not uniformly bounded in any neighborhood in $\mathbb{C}$. ●

In Chapter 2, we showed that every bounded sequence of complex numbers contains a convergent subsequence. Our goal in this section is to obtain analogous results for sequences of functions. It is not clear, at this point, what form of convergence is most reasonable or most applicable. To help clarify the situation, we need the following definition. A family $\mathcal{F}$ of functions is said to be *normal* in a domain D if every sequence $\{f_n\}$ in $\mathcal{F}$ contains a subsequence $\{f_{n_k}\}$ that converges uniformly on each compact subset of D.

As an example, the family consisting of the sequence $\{z^n\}$ is normal in the domain $|z| < 1$. In fact, the sequence itself converges uniformly to zero on every compact subset of $|z| < 1$. Note, however, that neither the sequence nor any subsequence converges uniformly in the whole domain.

Just as a bounded sequence may contain different subsequences that converge to different limits, so may a normal family contain different sequences that converge uniformly on compact subsets to different functions. To illustrate, set

$$f_n(z) = \begin{cases} z^n & \text{if } n \text{ odd,} \\ 1 - z^n & \text{if } n \text{ even.} \end{cases}$$

Then $\{f_{2n+1}\}$ converges uniformly to 0 and $\{f_{2n}\}$ converges uniformly to 1 on all compact subsets of $|z| < 1$.

A set of points E is said to be *dense* in a set A if every neighborhood of each point in A contains points of E. Every domain in the plane contains a dense sequence of points (for example, the set of points in the domain having both coordinates rational is countable, and so may be expressed as a sequence). Before proving the major result of this section, we need the following:

Lemma 11.13. *Suppose $\{f_n(z)\}$ is a sequence of analytic functions that is locally uniformly bounded in a domain D. If $\{f_n(z)\}$ converges at all points of a dense subset of D, then it converges uniformly on each compact subset of D.*

Proof. Given a compact set K contained in D, we wish to show that the sequence $\{f_n(z)\}$ converges uniformly on K. By Theorem 11.11, $\{f_n(z)\}$ is equicontinuous on K. Thus to each $\epsilon > 0$, there corresponds a $\delta > 0$ such that

$$|f_n(z) - f_n(z')| < \epsilon/3 \quad \text{for } |z - z'| < \delta, \tag{11.7}$$

where z, z' are any points in K and n is arbitrary. Since K is compact, finitely many, say p, neighborhoods of radius $\delta/2$ cover K. In each of these p neighborhoods, choose a point z_k $(k = 1, 2, \ldots, p)$ from the dense subset of K, at which $\{f_n\}$ converges. Next choose n and m large enough so that

$$|f_n(z_k) - f_m(z_k)| < \epsilon/3 \quad \text{for } k = 1, 2, \dots, p. \tag{11.8}$$

In view of (11.7) and (11.8), we see that, to each z in K, there corresponds a z_k in K such that

$$\begin{aligned} |f_n(z) - f_m(z)| &\leq |f_n(z) - f_n(z_k)| + |f_n(z_k) - f_m(z_k)| + |f_m(z_k) - f_m(z)| \\ &< \epsilon. \end{aligned}$$

Hence the sequence $\{f_n(z)\}$ is uniformly Cauchy on K, and must therefore converge uniformly on K. ■

Note the lemma concludes that $\{f_n(z)\}$ is a normal family in D. We will now show, by a diagonalization process, that this conclusion is true without the assumption that the sequence converges on a dense subset.

Theorem 11.14. (Montel's Theorem) *If $\mathcal{F}$ is a locally uniformly bounded family of analytic functions in a domain D, then $\mathcal{F}$ is a normal family in D.*

Proof. Given a sequence $\{f_n\}$ of functions in $\mathcal{F}$, we must show that some subsequence of $\{f_n\}$ converges uniformly on compact subsets. Choose any sequence of points $\{z_k\}$ that is dense in D. According to Lemma 11.13, it suffices to construct a subsequence of $\{f_n\}$ that converges at each point of the sequence $\{z_k\}$. By hypothesis, the sequence $\{f_n(z_1)\}$ of complex numbers is bounded. Hence by the Bolzano–Weierstrass property (see Theorem 2.17), there exists a subsequence of $\{f_n\}$, which we shall denote by $\{f_{n,1}\}$, that converges at z_1. But the sequence of $\{f_{n,1}(z_2)\}$ of points is also bounded. Thus there is a subsequence $\{f_{n,2}\}$ of $\{f_{n,1}\}$ that converges at z_2. Since $\{f_{n,2}\}$ is a subsequence of $\{f_{n,1}\}$, it must also converge at z_1.

Continuing the process, for each positive integer m, we obtain the mth subsequence $\{f_{n,m}\}$ of $\{f_n\}$ so that it converges at $z_1, z_2, \dots, z_m$. As seen in the chart below,

$$\begin{array}{cccccc} f_{1,1}(z), & f_{2,1}(z), & f_{3,1}(z), & \dots & f_{m,1}(z), & \dots \\ f_{1,2}(z), & f_{2,2}(z), & f_{3,2}(z), & \dots & f_{m,2}(z), & \dots \\ f_{1,3}(z), & f_{2,3}(z), & f_{3,3}(z), & \dots & f_{m,3}(z), & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ f_{1,m}(z), & f_{2,m}(z), & f_{3,m}(z), & \dots & f_{m,m}(z), & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{array}$$

the mth sequence of complex functions converges at z_m and all preceding points of the sequence $\{z_k\}$. Now consider the sequence $\{f_{n,n}(z)\}$, which makes up the diagonal of the chart. For each fixed m, the sequence $\{f_{n,n}(z_m)\}$, $n \geq m$, is a subsequence of the convergent sequence $\{f_{n,m}(z_m)\}$, and hence converges. Therefore, $\{f_{n,n}(z)\}$ is a subsequence of $\{f_n\}$ that converges at all points of the sequence $\{z_k\}$. This completes the proof. ■

The Bolzano–Weierstrass theorem guarantees the existence of a limit point for every bounded infinite set of points, and consequently the existence of a convergent subsequence for every bounded sequence. Montel's theorem can be viewed as an "analytic function" analog to Bolzano's theorem. It guarantees, in some sense, the existence of a convergent sequence of functions associated with every locally uniformly bounded family of analytic functions.

Carrying the analogy one step further, both theorems suffer from the same deficiency. The limit point of Bolzano's theorem need not be a member of the set, while the convergent function of Montel's need not be a member of the normal family. For example, the sequence $\{z^n\}$ is a normal family in $|z| < 1$ because it converges uniformly to 0 on all compact subsets of $|z| < 1$. However, 0 is not a member of the family $\{z^n\}$.

Recall that a bounded set that contains all its limit points is compact. This leads to the following definition. A normal family $\mathcal{F}$ of functions is said to be *compact* if the uniform limits of all sequences converging in $\mathcal{F}$ are themselves members of $\mathcal{F}$.

Example 11.15. The family $\mathcal{F}$ of functions of the form

$$f(z) = 1 + \sum_{n=1}^{\infty} a_n z^n$$

that are analytic with positive real part in the disk $|z| < 1$ is a compact, normal family. By Theorem 10.42, all functions $f \in \mathcal{F}$ satisfy the inequality

$$|f(z)| \le \frac{1+|z|}{1-|z|} \quad (|z| = r < 1).$$

Hence $\mathcal{F}$ is locally uniformly bounded and, by Montel's theorem, is normal. To show compactness, suppose a sequence $\{f_n\}$ of functions in $\mathcal{F}$ converges uniformly to a function g. We wish to show that $g \in \mathcal{F}$. By Theorem 8.16, g is analytic in $|z| < 1$. Since $f_n(0) = 1$ for every n, $g(0) = 1$. Since $\operatorname{Re} f_n(z) > 0$ for every n, $\operatorname{Re} g(z) \ge 0$ for $|z| < 1$. But then by the open mapping theorem, we must have $\operatorname{Re} g(z) > 0$ for $|z| < 1$. Thus $g \in \mathcal{F}$, and $\mathcal{F}$ is compact. ●

Questions 11.16.

1. What kinds of families of functions are locally uniformly bounded but not uniformly bounded?
2. Is the family of polynomials locally uniformly bounded on some set?
3. If $\mathcal{F}$ is a uniformly bounded family of analytic functions, is $\mathcal{F}^{(n)}$ also uniformly bounded?
4. If a family of functions is uniformly bounded at each point in a domain, is the family locally uniformly bounded?
5. Where, in the proof of Theorem 11.7, did we use the fact that the set K was a disk?

6. What is an important distinction between a dense sequence and a dense set?
7. What kinds of normal families have more than one subsequential limit function?
8. Can a normal family have infinitely many subsequential limit functions?

Exercises 11.17.

1. Suppose that for each point in a domain D there corresponds a neighborhood in which a family $\mathcal{F}$ is equicontinuous. Show that $\mathcal{F}$ is equicontinuous on compact subsets of D. Is $\mathcal{F}$ equicontinuous in D?
2. Show that the sequence $\{nz\}$ is not equicontinuous in any region.
3. If $\mathcal{F}$ is locally uniformly bounded family of analytic functions in a domain D, show that $\mathcal{F}'$, the family of functions consisting of the derivatives of functions in $\mathcal{F}$, is equicontinuous on compact subsets of D.
4. Suppose $\mathcal{F}$ is a normal family of analytic functions in the disk $|z| < 1$. Let $\mathcal{G}$ be the family of functions of the form $g(z) = \int_0^z f(\zeta)\, d\zeta$, where $f \in \mathcal{F}$. Show that $\mathcal{G}$ is normal in $|z| < 1$.
5. Show that the sequence $\{f_n(z)\}$ defined by
$$f_n(z) = \begin{cases} z^n & \text{if } n \text{ odd} \\ 1 - z^n & \text{if } n \text{ even,} \end{cases}$$
forms a normal family in the disk $|z| < 1$.
6. Show that the family of functions of the form $f(z) = \sum_{n=0}^{\infty} a_n z^n$, where $|a_n| \le n$, is a compact normal family of analytic functions in the disk $|z| < 1$.
7. Let $\mathcal{F}$ be the family consisting of all functions $f(z)$ that are analytic in a domain D with $|f(z)| \le M$ in D. Show that $\mathcal{F}$ is a compact, normal family in D.

11.3 Riemann Mapping Theorem

We have already discussed a number of examples of analytic functions between various domains of the complex plane. In some cases, we have given complete characterizations for mappings between certain domains such as disks and half-planes. Also, we know from the open mapping theorem that nonconstant analytic functions map domains into domains. Now, suppose D_1 and D_2 are simply connected domains. Then there is almost always an analytic function mapping D_1 onto D_2. We first discuss a "typical" exception. Suppose D_1 is the whole plane and D_2 is the disk $|z| < 1$. There can be no function analytic in the plane (entire) that maps onto the (bounded) disk $|z| < 1$, for, according to Liouville's theorem, constant functions are the only entire functions whose images are contained in the disk. Our major theorem of this section says that a one-to-one analytic mapping exists between any two simply connected domains, neither of which is the whole plane. Before proving this remarkable (existence) result, we shall need some preliminaries concerning *univalent* (a fancy term for one-to-one) functions.

Theorem 11.18. *Suppose $\{f_n(z)\}$ is a sequence of analytic, univalent functions defined in a domain D and converging uniformly on each compact subset of D to a nonconstant function $f(z)$. Then $f(z)$ is analytic and univalent in D.*

Proof. The analyticity of f follows from Theorem 8.16. To prove the univalence of f, assume there are distinct points z_0, z_1 in D for which $f(z_0) = f(z_1) = a$. We can find $r > 0$ (e.g., $r < |z_0 - z_1|/2$) so small that the closed disks centered at z_0 and z_1 with radius r are mutually disjoint and are contained in D. Assume further that $f(z) \neq a$ on the circles $C_0 : |z - z_0| = r$ and $C_1 : |z - z_1| = r$. This is possible because f is nonconstant. Let

$$m = \min_{z \in C_0 \cup C_1} |f(z) - a|.$$

Now choose n sufficiently large so that $|f_n(z) - f(z)| < m$ on $C_0 \cup C_1$. So, on $C_0 \cup C_1$,

$$|f(z) - a)| > m > |f_n(z) - f(z)| \text{ for large } n.$$

By Rouche's theorem, the function

$$f_n(z) - a = (f_n(z) - f(z)) + (f(z) - a)$$

has at least one zero inside C_0 and at least one zero inside C_1. This contradicts the univalence of $f_n(z)$ in D. ■

Note that it is possible for the uniform limit of a sequence of univalent functions to be constant. For example, the univalent sequence $f_n(z) = z/n$ converges uniformly to $f(z) = 0$ on any compact subset of $\mathbb{C}$. Thus the uniform limit of a sequence of univalent functions need not be univalent.

Theorem 11.19. *Suppose $f(z)$ is analytic and univalent in a domain D, and that $g(z)$ is analytic and univalent on the image of D under $f(z)$. Then the composition function $g(f(z))$ is analytic and univalent in D.*

Proof. The analyticity of $g(f(z))$ follows from Theorem 5.6. To show univalence, suppose

$$g(f(z_0)) = g(f(z_1)) \text{ for } z_0, z_1 \in D.$$

By the univalence of g, we have $f(z_0) = f(z_1)$. From the univalence of f, $z_0 = z_1$ and the theorem is proved. ■

Theorem 11.20. *Suppose f, mapping a domain D_1 onto D_2, is analytic and univalent in D_1. Then the inverse function g, defined by $g(f(z)) = z$ for all $z \in D_1$, is an analytic and univalent mapping from D_2 onto D_1.*

Proof. The univalence of g is an immediate consequence of the univalence of f. To show analyticity, fix a point $w_0 \in D_2$. Then $w_0 = f(z_0)$ for a unique $z_0 \in D_1$. Setting $w = f(z)$, we have

$$\frac{g(w) - g(w_0)}{w - w_0} = \frac{z - z_0}{f(z) - f(z_0)}. \tag{11.9}$$

Since f maps open sets onto open sets (Theorem 9.55), g is continuous in D_2. Thus $z \to z_0$ as $w \to w_0$. By Theorem 11.3, $f'(z_0) \neq 0$. Hence we may take limits in (11.9) to obtain $g'(w_0) = g'(f(z_0)) = 1/(f'(z_0))$. Therefore g is analytic in D_2, and the theorem is proved. ■

If f and g are analytic and univalent in domains D_1 and D_2, respectively, and map onto the disk $|z| < 1$, then $g^{-1}(f(z))$ is an analytic and univalent mapping from D_1 onto D_2 (see Figure 11.9).

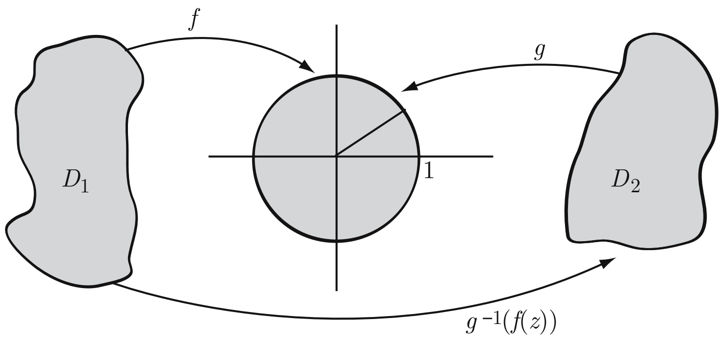


Figure 11.9.

Thus the set of domains that may be mapped analytically and univalently onto the interior of the unit disk can also be mapped analytically and univalently onto one another.

Suppose f is analytic and univalent in D and maps onto $|z| < 1$. Are there other functions with the same property? In general, there are infinitely many. To see this, recall from Section 3.2 (see Theorem 3.21) that all functions of the form

$$g(z) = e^{i\alpha} \frac{z - z_0}{1 - \overline{z}_0 z} \quad (|z_0| < 1,\ \alpha \ \text{ real}) \tag{11.10}$$

map the interior of the unit circle onto itself. Hence the functions $g(f(z))$ and $f(z)$ simultaneously map D onto $|z| < 1$. Our next result suggests conditions for establishing a unique mapping function.

Given a domain $D \subseteq \mathbb{C}$, we define the group of analytic automorphisms of D as follows: If $f : D \to D$ is an analytic function that is one-to-one and onto, then $f(z)$ is called an analytic/holomorphic automorphism of D. That is, $f(z)$

is called a conformal self-mapping of D. The set of all analytic automorphisms of D form what is called an "automorphism group" (with composition as the group operation) of D, and is denoted by $\mathrm{Aut}\,(D)$. The Schwarz lemma can be used to describe the automorphism groups of the upper half-plane, and the unit disk Δ (see also Theorems 3.18 and 3.21). It is easy to see that

$$\mathrm{Aut}_{\,a}(D) = \{f \in \mathrm{Aut}\,(D) : f(a) = a\}$$

forms a subgroup of the group $\mathrm{Aut}\,(D)$. Our next result is a reformulation of Theorem 3.21 in the language of automorphisms, but the new proof uses the Schwarz lemma.

Theorem 11.21. *We have*

$$\mathrm{Aut}\,(\Delta) = \left\{ e^{i\alpha}\left(\frac{z-a}{1-\overline{a}z}\right) : |a|<1,\ 0 \le \alpha \le 2\pi \right\}.$$

In particular, $\mathrm{Aut}_{\,0}(\Delta) := \{f \in \mathrm{Aut}\,(\Delta) : f(0) = 0\} = \{e^{i\alpha}z : \alpha \textit{ real}\}$.

Proof. Let $a \in \Delta$, and

$$\varphi_a(z) = \frac{a-z}{1-\overline{a}z}.$$

Obviously, φ_a is analytic for $|z| < 1/|a|$ ($|a| < 1$), $\varphi_a(\Delta) \subseteq \Delta$, and $\varphi_a(\partial\Delta) = \partial\Delta$. Moreover, φ_a is univalent on Δ and $(\varphi_a)^{-1} = \varphi_a$. Thus, $\varphi_a \in \mathrm{Aut}\,(\Delta)$. Also, the rotation $e^{i\theta}\varphi_a(z)$ ($\theta \in \mathbb{R}$) belongs to $\mathrm{Aut}\,(\Delta)$.

Conversely, let $f \in \mathrm{Aut}\,(\Delta)$. Then there exists a $b \in \Delta$ such that $f(0) = b$. Then $F(z)$ defined by $F = \varphi_b \circ f$ is also analytic and univalent in Δ, F maps Δ onto Δ, and $F(0) = 0$. By the Schwarz lemma,

$$|F(z)| \le |z| \ \text{ for } \ z \in \Delta.$$

Since F is analytic and one-to-one on Δ, F^{-1} exists on Δ. Moreover, F^{-1} is analytic and one-to-one on Δ with $F^{-1}(0) = 0$. We may again apply the Schwarz lemma to F^{-1} and obtain $|F^{-1}(w)| \le |w|$ for $w \in \Delta$. If we take $w = F(z)$, we get

$$|z| \le |F(z)| \ \text{ for } \ z \in \Delta.$$

Hence, $|F(z)| = |z|$, and so $F(z) = \lambda z$ with $|\lambda| = 1$, or

$$\varphi_b(f(z)) = \lambda z \text{ or } \ f(z) = \varphi_b(\lambda z).$$

The desired result follows. ∎

Our next result suggests conditions for establishing a unique mapping function.

Lemma 11.22. *Suppose $f(z)$ is analytic and univalent in $|z| < 1$ and maps the disk onto itself. If $f(0) = 0$ and $f'(0) > 0$, then $f(z) = z$.*

Proof. As $\operatorname{Aut}_0(\Delta) := \{f \in \operatorname{Aut}(\Delta) : f(0) = 0\} = \{e^{i\alpha}z : \alpha \text{ real}\}$ and $f'(0) > 0$, the result follows. ∎

Two domains D_1 and D_2 are said to be *conformally equivalent* if there is a bijective analytic function mapping D_1 onto D_2. Both the existence and method of finding it are two important components for conformal mappings. We start with a couple of examples illustrating conformal mappings between standard simply connected domains. It follows that conformally equivalent domains are homeomorphic but not the converse.

Example 11.23. We are interested in showing that the upper half disk $D = \{z : |z| < 1, \operatorname{Im} z > 0\}$ and the unit disk $\Delta = \{z : |z| < 1\}$ are conformally equivalent.

Step 1: We consider

$$w_1 = f_1(z) = \frac{1}{1-z}.$$

Then we know that f_1 transforms the unit disk Δ onto the right half-plane $\operatorname{Re} w_1 > 1/2$. Rewriting

$$w_1 = f_1(z) = \frac{1-\overline{z}}{|1-z|^2} = \frac{1-x+iy}{|1-z|^2},$$

we see that $\operatorname{Im} w_1 > 0$ iff $\operatorname{Im} z > 0$. Moreover, $z = 1$ is a pole of $f_1(z)$, the segment $[-1, 1]$ maps onto the half-line $[1/2, \infty)$ and the upper half circle $\{z : |z| = 1, \operatorname{Im} z > 0\}$ onto the half-line $\{w_1 : \operatorname{Re} w_1 = 1/2, \operatorname{Im} w_1 > 0\}$. Therefore, f_1 maps D onto $D_1 = \{w_1 : \operatorname{Re} w_1 > 1/2, \operatorname{Im} w_1 > 0\}$.

Step 2: The map $w_2 = f_2(w_1) = w_1 - 1/2$ maps the domain D_1 onto the first quadrant $D_2 = \{w_2 : \operatorname{Re} w_2 > 0, \operatorname{Im} w_2 > 0\}$.

Step 3: The map $w_3 = f_3(w_2) = w_2^2$ maps D_2 onto the upper half-plane $\mathbb{H}^+ = \{w_3 : \operatorname{Im} w_3 > 0\}$.

Step 4: The map $w = f_4(w_3) = \frac{w_3 - i}{w_3 + i}$ carries the upper half-plane $\mathbb{H}^+$ onto the unit disk $\{w : |w| < 1\}$. Finally a map f with the desired property is a composition

$$w = f(z) = (f_4 \circ f_3 \circ f_2 \circ f_1)(z) = f_4(f_3(f_2(f_1(z))))$$

which gives

$$w = f(z) = \frac{(1+z)^2 - 4i(1-z)^2}{(1+z)^2 + 4i(1-z)^2}.$$ •

Example 11.24. Let $D = \{z : |z| < 1, |z - 1/2| > 1/2\}$. Now we want to find a conformal map of D onto the unit disk Δ. As we can see from the picture, it suffices to focus on certain key points to understand the sequence of mappings considered here. If $w_1 = 1/(1-z)$, then $z = 1 - 1/w_1$ and

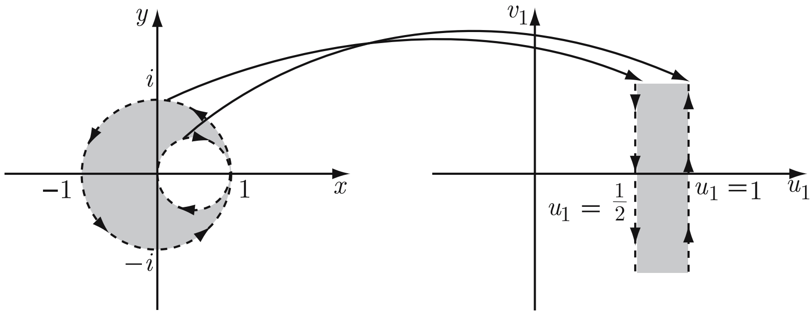


Figure 11.10. A conformal map of D onto the strip

$$\begin{cases} |z| < 1 & \Longleftrightarrow \operatorname{Re} w_1 > 1/2 \\ |z - 1/2| > 1/2 & \Longleftrightarrow \operatorname{Re} w_1 < 1. \end{cases}$$

Because of the basic property of Möbius transformations, it follows easily that f_1 maps D onto the strip $D_1 = \{w_1 : 1/2 < \operatorname{Re} w_1 < 1\}$. A similar explanation may be provided for other mappings. Finally, the composition

$$w = f(z) = f_4 \circ f_3 \circ f_2 \circ f_1(z)$$

gives the formula which does the required job, where

$$w_2 = f_2(w_1) = i\pi(w_1 - 1/2),\ w_3 = f_3(w_2) = e^{w_2},\ f_4(w_3) = \frac{w_3 - i}{w_3 + i}. \qquad \bullet$$

We are now ready to formally state and prove the Riemann mapping theorem which is a classical example of existence theorems.

Theorem 11.25. (Riemann Mapping Theorem) *Suppose D is a simply connected domain, other than the whole plane, and z_0 is a point in D. Then there exists a unique function $f(z)$, analytic and univalent in D, which maps D onto the disk $|w| < 1$ in such a manner that $f(z_0) = 0$ and $f'(z_0) > 0$.*

Proof. We first prove the uniqueness of the mapping function f. If g_1 and g_2 are two functions each of which maps D onto the unit disk $|w| < 1$ in the prescribed manner, then $h = g_2 \circ g_1^{-1}$ is an analytic and univalent mapping of the unit disk $|w| < 1$ onto itself. Furthermore,

$$h(0) = g_2(g_1^{-1}(0)) = g_2(z_0) = 0$$

and, because $g_1'(z_0) > 0$ and $g_2'(z_0) > 0$,

$$h'(0) = g_2'(g_1^{-1}(0))(g_1^{-1})'(0) = \frac{g_2'(z_0)}{g_1'(z_0)} > 0.$$

Hence, by Lemma 11.22, h is the identity function. That is, $g_1(z) = g_2(z)$ and uniqueness is proved.

To prove existence of the mapping function, we first show that there is an analytic and univalent function mapping D *into* the disk $|w| < 1$. Since D is not the whole plane $\mathbb{C}$, there is a point $a \in \mathbb{C}\setminus D$. If there is actually a disk $|z-a| < \epsilon$ outside of D, then $|z-a| > \epsilon$ for all points z in D. In this case,

$$w = \frac{\epsilon}{z-a}$$

is an analytic and univalent function that maps all points of D into the unit disk $|w| < 1$. Thus, the proof follows if D is a bounded domain. However, if D is unbounded, then it is possible that the complement of D does not contain any disk. For instance, D might be the plane minus a ray from some point z_0 to ∞. This kind of difficulty will be avoided by considering a branch of the square root function, which maps a domain onto one "half" its size.

According to Corollary 7.52, if $a \in \mathbb{C}\setminus D$, then there exists an analytic function $\phi : D \to \mathbb{C}$, called analytic branch of $(z-a)^{1/2}$ with $\phi^2(z) = z-a$ so that $\phi(z) = \sqrt{z-a}$. Furthermore, $\phi(z)$ is univalent in D. For if $\phi(z_1) = \phi(z_2)$ for $z_1, z_2 \in D$, then

$$[\phi(z_1)]^2 = [\phi(z_2)]^2, \quad \text{i.e.,} \quad z_1 - a = z_2 - a.$$

Now let $D' = \phi(D)$. Then D' is simply connected since D is simply connected. Then the complement of D' contains a disk. To see this, we will show that points b and $-b$ cannot simultaneously be in D'. For if they are, then there exist two points z_1 and z_2 in D such that $\phi(z_1) = b$ and $\phi(z_2) = -b$. Now,

$$\begin{aligned}
\phi(z_1) = -\phi(z_2) &\Longrightarrow [\phi(z_1)]^2 = [\phi(z_2)]^2 \\
&\Longrightarrow z_1 - a = z_2 - a, \quad \text{i.e.,} \quad z_1 = z_2 \\
&\Longrightarrow b = -b, \quad \text{i.e.,} \quad \phi(z_1) = 0 = \phi(z_2) \\
&\Longrightarrow z_2 = a \in \mathbb{C}\setminus D,
\end{aligned}$$

contradicting the fact that z_1 and z_2 are distinct.

Next choose a point $w_0 \in D'$ and an $\epsilon > 0$ so that the disk $|w - w_0| < \epsilon$ is contained in D'. Then the disk $|w + w_0| < \epsilon$ is contained in the complement $\mathbb{C}\setminus D'$. Hence the function

$$\psi(w) = \frac{\epsilon}{w + w_0}$$

maps D' into the unit disk, because $|w + w_0| > \epsilon$ for all $w \in D'$. Therefore, the composition

$$f(z) = \psi(\phi(z)) = \frac{\epsilon}{\phi(z) + w_0}$$

is analytic and univalent in D and maps D into the unit disk. By a suitable bilinear transformation (fill in details!), we can transform this function into a function $f_0(z)$ satisfying the additional conditions $f_0(z_0) = 0$ and $f_0'(z_0) > 0$.

Let $\mathcal{F}$ denote the family of all analytic functions $g : D \to \mathbb{C}$ such that $g(z)$ is univalent in D, $g(z_0) = 0$, $g'(z_0) > 0$, and satisfies $|g(z)| < 1$ for all z in D. The family $\mathcal{F}$ is nonempty because $f_0(z) \in \mathcal{F}$. Certainly the function whose existence we are determined to prove must also be in the family $\mathcal{F}$. It will be shown that the desired function has a larger derivative at z_0 than any other function in $\mathcal{F}$. To show the existence of a function in $\mathcal{F}$ with a maximum derivative at z_0, we will rely on the theory of normal families.

Since the family $\mathcal{F}$ is locally uniformly bounded (in fact, uniformly bounded) in D, it follows from Theorem 11.14 that $\mathcal{F}$ is a normal family. Set

$$A = \text{lub}\,\{g'(z_0) : g \in \mathcal{F}\}.$$

Then, $A > 0$ because $g'(z_0) > 0$ for each $g \in \mathcal{F}$. But A may be infinite. By the definition of A, there is a sequence $\{f_n\}$ of functions in $\mathcal{F}$ such that $f_n'(z_0) \to A$. By the normality of $\mathcal{F}$, there exists a subsequence $\{f_{n_k}\}$ that converges uniformly on the compact subsets of D to an analytic function $f(z)$. An application of Corollary 8.18 shows that $f'(z_0) = A$, so that A is finite. Since $f'(z_0) \geq f_0'(z_0) > 0$, the function $f(z)$ is not constant in D. It thus follows from Theorem 11.18 that $f(z)$ is univalent and, consequently, a member of $\mathcal{F}$.

We shall now show that this f maps D onto the unit disk, and so it is the required function. For the sake of obtaining a contradiction we suppose that $f(D)$ is not the whole unit disk $|w| < 1$. Then $f(z) \neq \alpha$ for some α with $|\alpha| < 1$. By the definition of analytic branch of square roots, there exists an analytic function $F(z)$ in D so that

$$F(z)^2 = \frac{f(z) - \alpha}{1 - \overline{\alpha} f(z)}.$$

The univalence of $F(z)$ follows from the univalence of $f(z)$, and the inequality $|F(z)| < 1$ follows from the inequality $|f(z)| < 1$. However, $F(z)$ is not properly normalized. We therefore consider the function

$$G(z) = \frac{|F'(z_0)|}{F'(z_0)} \frac{F(z) - F(z_0)}{1 - \overline{F(z_0)} F(z)},$$

which satisfies $G(z_0) = 0$ and $G'(z_0) > 0$, so that $G(z) \in \mathcal{F}$. Moreover,

$$G'(z_0) = \frac{|F'(z_0)|}{1 - |F(z_0)|^2} = \frac{1 + |\alpha|}{2\sqrt{|\alpha|}} A > A = f'(z_0),$$

contradicting the maximality of $f'(z_0)$. Thus $f(z)$ omits no values inside the unit disk, and the proof is complete. ■

Remark 11.26. Since univalence in a domain guarantees a nonvanishing derivative, the Riemann mapping theorem shows that any two simply connected domains (neither of which is the plane) are conformally equivalent.

In the proof of Theorem 11.25, we assumed that an analytic, univalent function maps simply connected domains onto simply connected domains. In elementary topology, it is proved that the one-to-one continuous image of a simply connected domain cannot be multiply connected. Thus, we conclude that no simply connected domain can be conformally equivalent to a multiply connected domain. ●

Remark 11.27. Recall that a bilinear transformation maps circles and straight lines onto circles and straight lines. Hence any conformal mapping of a domain, other than a disk or a half-plane, onto the interior of the unit circle must be accomplished by a function other than a bilinear transformation. Furthermore, by the uniqueness property of the Riemann mapping theorem, no univalent function other than a bilinear transformation can map a disk or a half-plane onto the interior of the unit circle.

At this point, we must reflect on a sobering thought. The Riemann mapping theorem, like many existence theorems, has the drawback of not furnishing much insight into the actual construction. Therefore, given two "unfamiliar" simply connected domains, we must plod along as before to develop techniques for determining an appropriate mapping function. ●

Remark 11.28. The mapping of the interior of an arbitrary polygon onto the interior of the unit circle, whose existence is guaranteed by the theorem, can be found explicitly. This is accomplished in several stages. The *Schwarz–Christoffel formula* gives an analytic and univalent mapping of the upper half-plane onto the interior of an arbitrary polygon. For a complete discussion of the Schwarz–Christoffel transformation, we refer the reader to Nehari [N]. Composing the inverse of such a mapping with a bilinear transformation from the upper half-plane onto the open unit disk (see Section 3.3) gives the desired mapping. ●

Example 11.29. Let $f : \Omega \to \Omega$ be analytic in a simply connected domain Ω ($\neq \mathbb{C}$) having a fixed point in Ω. Then it can easily be shown that $|f'(a)| \leq 1$, and if $|f'(a)| = 1$, then f is actually a homeomorphism from Ω onto Ω.

The Riemann mapping theorem assures the existence of a bijective conformal map $\phi : \Omega \to \Delta$ such that $\phi(a) = 0$. Then we see that g defined by

$$g(z) = \phi \circ f \circ \phi^{-1}(z)$$

maps Δ into Δ and satisfies the hypothesis of the Schwarz lemma. Now, we easily see that $g'(0) = f'(a)$ and so $|f'(a)| \leq 1$, because $|g'(0)| \leq 1$. Moreover,

$$\begin{aligned} |f'(a)| = 1 &\Longrightarrow |g'(0)| = 1 \\ &\Longrightarrow g(z) = e^{i\alpha} z \quad \text{(by the Schwarz lemma)} \\ &\Longrightarrow \phi \circ f \circ \phi^{-1}(z) = e^{i\alpha} z \\ &\Longrightarrow f(z) = \phi^{-1}(e^{i\alpha}\phi(z)) \end{aligned}$$

which implies that f must be a bijective mapping from Ω onto Ω, because $\phi : \Omega \to \Delta$ and $\phi^{-1} : \Delta \to \Omega$ are bijective maps. ●

Questions 11.30.

1. Must the convergence be uniform in Theorem 11.18 in order for the conclusion to be valid?
2. Are there conformal mappings from multiply connected domains onto multiply connected domains?
3. If $f(z)$ is analytic and conformal in a domain D_1 and maps D_1 onto D_2, are D_1 and D_2 conformally equivalent?
4. What other initial conditions could we have prescribed in the Riemann mapping theorem to guarantee uniqueness?
5. Does there exist a one-to-one conformal mapping from the unit disk onto the disk minus the origin?
6. If two domains are conformally equivalent, what can be said about their boundaries?
7. Does there always exist an analytic function which maps a simply connected domain $\Omega(\neq \mathbb{C})$ into the unit disk $|z| < 1$?
8. Let $\Omega\,(\neq \mathbb{C})$ be a simply connected domain and let $\mathcal{F}$ be the set of all one-to-one analytic functions which map Ω into the unit disk $|z| < 1$, and $a \in \Omega$. If $f \in \mathcal{F}$ and is not onto, is there a function $g \in \mathcal{F}$ such that $|g'(a)| > |f'(a)|$?
9. Are the plane $\mathbb{C}$ and the unit disk $|z| < 1$ conformally equivalent? Are they homeomorphic?
10. Are the plane $\mathbb{C}$ and the upper half-plane $\operatorname{Im} w > 0$ conformally equivalent? Are they homeomorphic?
11. In the statement of the Riemann mapping theorem, why do we require the domain D to be a proper subset of $\mathbb{C}$? Does the theorem still hold if we remove that assumption?
12. Does the proof of the Riemann mapping theorem use the fact that every nonvanishing analytic function in a simply connected domain D admits analytic square root function in D?
13. Where, in the proof of the Riemann mapping theorem, did we require the domain to be simply connected?
14. Why was it necessary to first show that some function mapped the domain *into* the unit disk?
15. Why does the function $G(z)$, constructed in the proof of the Riemann mapping theorem, work?
16. What is a conformal map between the upper half-plane $\mathbb{H}^+ = \{z : \operatorname{Im} z > 0\}$ and $\mathbb{C} \setminus [0, \infty)$?
17. What is a conformal map between the right half-plane $D_1 = \{z : \operatorname{Re} z > 0\}$ and $D_2 = \{z : |\operatorname{Arg} z| < \pi/8\}$?
18. What is a conformal map between the strip $D_1 = \{z : 0 < \operatorname{Im} z < \pi/2\}$ and the upper half-plane $\mathbb{H}^+$?

19. What is a conformal map between the strip $D_1 = \{z : 0 < \operatorname{Im} z < \alpha\}$ and the upper half-plane $\mathbb{H}^+$?
20. What is a conformal map between the infinite strip $|\operatorname{Re} z| < \pi/2$ and the unit disk $|w| < 1$?
21. What is a conformal map between the unit disk $|z| < 1$ and $\mathbb{C} \setminus \Delta$?

Exercises 11.31.

1. Suppose $f(z)$ is analytic at z_0 with $f'(z_0) \neq 0$. Show that there exist neighborhoods U and V of z_0 and $f(z_0)$, respectively, such that $f(z)$ is a univalent mapping from U onto V.
2. Show that the plane is not conformally equivalent to the upper half-plane. More generally, show that the plane is only conformally equivalent to itself.
3. Let $D_1 = \{z : 0 < \operatorname{Re} z,\ \operatorname{Im} z < \infty\}$ and $D_2 = \{w : \operatorname{Im} w > 0\}$ be the open first quadrant and the upper half-plane, respectively. By the Riemann mapping theorem D_1 and D_2 are conformally equivalent. Show that $f(z) = z^2$ does this job.
4. Let $D_1 = \{z : |\operatorname{Re} z| < \pi/2\}$ and $D_2 = \{w : \operatorname{Re} w > 0\}$. Show that $f : D_1 \to D_2$ given by $f(z) = e^{iz}$ is conformal.
5. Even though the interior of a square can be mapped conformally onto the interior of a circle, show that no square can be mapped conformally onto a circle.
6. Let D_1 be the annulus $0 < r_1 < |z| < R_1$ and D_2 be the annulus $0 < r_2 < |z| < R_2$. If
$$\frac{R_1}{r_1} = \frac{R_2}{r_2},$$
construct an analytic and univalent function that maps D_1 onto D_2.
7. Suppose D_1 and D_2 are conformally equivalent, and that D_2 and D_3 are conformally equivalent. Show that D_1 and D_3 are conformally equivalent.

11.4 The Class $\mathcal{S}$

We continue our investigation of univalent functions—a specialized topics in complex analysis. Analytically, a univalent function has a nonvanishing derivative (Theorem 11.3); geometrically, a univalent function maps simple curves onto simple curves.

Functions that are both analytic and univalent have a nice property of mapping simply connected domains onto simply connected domains. By the Riemann mapping theorem, we can associate a univalent function defined in an arbitrary simply connected domain (other than the whole plane) with one defined in the unit disk. Therefore, we shall restrict the domain on which these functions are defined to the disk $|z| < 1$. Our results will have a nicer form if we also assume that the function has a zero (hence its only zero) at the origin and that its derivative is equal to one at the origin. Since the derivative of a univalent function never vanishes, every univalent function $h(z)$ may be reduced to a function of this form by replacing it with

$$f(z) = \frac{h(z) - h(0)}{h'(0)}.$$

We shall denote by $\mathcal{S}$ the class of all functions $f(z)$ that are analytic and univalent in the unit disk $|z| < 1$, and are normalized by the conditions $f(0) = 0$ and $f'(0) = 1$. Thus a function $f(z)$ in $\mathcal{S}$ has the power series representation

$$f(z) = z + a_2 z^2 + a_3 z^3 + \cdots \qquad (|z| < 1).$$

We shall denote by $\mathcal{T}$ the class of all functions of the form

$$g(z) = z + b_0 + \frac{b_1}{z} + \frac{b_2}{z^2} + \cdots$$

that are analytic and univalent in the domain $|z| > 1$. The following relationship will enable us to deduce information about $\mathcal{S}$ from information about $\mathcal{T}$.

Theorem 11.32. *If $f(z) \in \mathcal{S}$, then $1/f(1/z) \in \mathcal{T}$.*

Proof. First suppose $1/f(1/z_1) = 1/f(1/z_2)$ $(|z_1| > 1, |z_2| > 1)$. Then $f(1/z_1) = f(1/z_2)$, where $|1/z_1| < 1$ and $|1/z_2| < 1$. The univalence of $1/f(1/z)$ $(|z| > 1)$ now follows from the univalence of $f(z)$ $(|z| < 1)$. The analyticity of $1/f(1/z)$ will be a consequence of the analyticity of $f(z)$ if we can show that $f(1/z) \neq 0$ for $|z| > 1$. If $f(1/z_0) = 0$ for $0 < |1/z_0| < 1$, then $f(0) = f(1/z_0) = 0$, contradicting the univalence of $f(z)$ for $|z| < 1$. Hence $1/f(1/z) \in \mathcal{T}$, and the proof is complete. ∎

The next theorem, because of its proof rather than its statement, is known as the *area theorem.*

Theorem 11.33. *If $g(z) = z + b_0 + (b_1/z) + (b_2/z^2) + \cdots$ is in $\mathcal{T}$, then $\sum_{n=1}^{\infty} n|b_n|^2 \leq 1$.*

Proof. The univalent function $g(z)$ maps the circle $|z| = r > 1$ onto a simple closed contour C. Set $g(z) = u(z) + iv(z)$. The area of the region R enclosed by C, denoted by $A(r)$, is

$$A(r) = \iint_R du\, dv.$$

Note that $A(r) > 0$ for each $r > 1$. If we now let $P(u, v) = -v/2$ and $Q(u, v) = u/2$, an application of Green's theorem yields

$$A(r) = \frac{1}{2}\int_C u\,dv - v\,du = \frac{1}{2}\int_0^{2\pi} \left(u\frac{\partial v}{\partial \theta} - v\frac{\partial u}{\partial \theta}\right) d\theta, \tag{11.11}$$

where $A(r) > 0$. By Exercise 5.2(13), we have $g'(z) = (1/iz)(\partial g/\partial \theta)$. To evaluate the line integral of (11.11), consider the integral

$$\begin{aligned} \frac{1}{2}\int_{|z|=r} \overline{g(z)}g'(z)\,dz &= \frac{1}{2}\int_0^{2\pi} (u - iv)\left[\frac{1}{iz}\left(\frac{\partial u}{\partial \theta} + i\frac{\partial v}{\partial \theta}\right)\right] iz\,d\theta \\ &= \frac{1}{2}\int_0^{2\pi} \left(u\frac{\partial u}{\partial \theta} + v\frac{\partial v}{\partial \theta}\right) d\theta + \frac{i}{2}\int_0^{2\pi} \left(u\frac{\partial v}{\partial \theta} - v\frac{\partial u}{\partial \theta}\right) d\theta, \end{aligned} \tag{11.12}$$

whose imaginary part corresponds to $A(r)$. In order to simplify (11.12), we write

$$\int_{|z|=r} \overline{g(z)}g'(z)\,dz = \int_{|z|=r} \left(\overline{z} + \sum_{m=0}^{\infty} \overline{b}_m (\overline{z})^{-m}\right)\left(1 - \sum_{n=1}^{\infty} nb_n z^{-n-1}\right) dz,$$

and note that

$$\int_{|z|=r} (\overline{z})^{-m} z^{-n-1}\,dz = \begin{cases} 2\pi i r^{-2m} & \text{if } n = m, \\ 0 & \text{if } n \neq m. \end{cases}$$

This leads to the identity

$$\begin{aligned} \frac{1}{2}\int_{|z|=r} \overline{g(z)}g'(z)\,dz &= \frac{1}{2}\int_{|z|=r} \overline{z}\,dz - \frac{1}{2}\int_{|z|=r} \frac{\sum_{n=1}^{\infty} n|b_n|^2 r^{-2n}}{z}\,dz \\ &= \pi i\left(r^2 - \sum_{n=1}^{\infty} \frac{n|b_n|^2}{r^{2n}}\right). \end{aligned}$$

Therefore (11.12) is purely imaginary, and

$$A(r) = \frac{1}{2}\int_0^{2\pi} \left(u\frac{\partial v}{\partial \theta} - v\frac{\partial u}{\partial \theta}\right) d\theta = \pi\left(r^2 - \sum_{n=1}^{\infty} \frac{n|b_n|^2}{r^{2n}}\right). \tag{11.13}$$

Since $A(r) > 0$, we have

$$r^2 - \sum_{n=1}^{\infty} \frac{n|b_n|^2}{r^{2n}} > 0 \quad (r > 1). \tag{11.14}$$

But (11.14) is valid for every $r > 1$ so that the result follows upon letting $r \to 1^+$. ■

Remark 11.34. According to (11.13), the area enclosed by the image of the circle $|z| = r$ is at most πr^2 (the area enclosed by the circle), with equality only for $g(z) = z + b_0$. Furthermore, equality in the conclusion of the theorem holds if and only if the area enclosed by the image of $|z| = r > 1$ becomes arbitrarily small as $r \to 1$. ●

Remark 11.35. If $b_1 = 1$, then $b_n = 0$ for $n > 1$. Recall that the properties of $g(z) = z + 1/z$ were extensively studied in Section 3.3. In particular, this function was shown to map $|z| = r > 1$ onto an ellipse, and the ellipse approaches the linear segment $[-2, 2]$ as r approaches 1. ●

The coefficient bound for functions in $\mathcal{T}$, as expressed by the area theorem, will enable us to obtain a coefficient bound for functions in $\mathcal{S}$. But first we need the following:

Lemma 11.36. *If $f(z) \in \mathcal{S}$, then $z\sqrt{f(z^2)/z^2} \in \mathcal{S}$.*

Proof. Set $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$. Then

$$f(z^2) = z^2[1 + a_2 z^2 + a_3 z^4 + \cdots] := z^2 h(z),$$

where $h(z)$ is analytic and never vanishes in the unit disk. Therefore, choosing a branch of $(h(z))^{1/2}$ with $(h(0))^{1/2} = 1$, we see that $g(z)$ defined by

$$g(z) = z\sqrt{\frac{f(z^2)}{z^2}} = z\sqrt{1 + a_2 z^2 + a_3 z^4 + \cdots} \tag{11.15}$$

is analytic with $g(0) = 0$ and $g'(0) = 1$. To prove that $g(z)$ is univalent, suppose $g(z_1) = g(z_2)$. Then $f(z_1^2) = f({z_2}^2)$, and the univalence of $f(z)$ shows that $z_1^2 = z_2^2$, that is, $z_1 = \pm z_2$. But from (11.15), we see that $g(z)$ is an odd function. Hence, $z_1 = -z_2$ implies $g(z_1) = -g(z_2)$, which is a contradiction unless $z_1 = z_2 = 0$. Therefore $z_1 = z_2$, thus establishing the univalence of $g(z)$. ■

Remark 11.37. It was necessary to write $z\sqrt{f(z^2)/z^2}$ instead of $\sqrt{f(z^2)}$ because $f(z^2)$ has a zero at the origin, which makes the expression

$$\sqrt{f(z^2)} = e^{(1/2)\,\mathrm{Log}\, f(z^2)}$$

meaningless. ●

Theorem 11.38. *If $f(z) = z + a_2 z^2 + \cdots$ is in $\mathcal{S}$, then $|a_2| \le 2$.*

Proof. By Lemma 11.36, $g(z) = z\sqrt{f(z^2)/z^2} \in \mathcal{S}$. We can verify from the expansion in (11.15) that $g'''(0) = 3a_2$. Thus we may write

$$g(z) = z + \frac{a_2}{2} z^3 + \cdots .$$

In view of Theorem 11.32, the Laurent expansion for $1/g(1/z)$ shows that

$$\frac{1}{g(1/z)} = \frac{1}{(1/z)[1 + (a_2/2)z^2 + \cdots]} = z - \frac{a_2}{2}\frac{1}{z} + \cdots \in \mathcal{T}.$$

Applying Theorem 11.33, we find that $|a_2/2|^2 \le 1$, i.e., $|a_2| \le 2$. ■

Remark 11.39. Retracing the steps in the proof, we can determine when equality holds. For if $a_2 = 2e^{i\alpha}$, α real, then $1/g(1/z) = z - e^{i\alpha}/z$. But this means that $g(z) = z/(1-e^{i\alpha}z)^2 = z\sqrt{f(z^2)/z^2}$, so that

$$f(z) = \frac{z}{(1-e^{i\alpha}z)^2} = z + 2e^{i\alpha}z^2 + 3e^{2i\alpha}z^3 + \cdots . \tag{11.16}$$

For each $\alpha \in \mathbb{R}$, this function is known as the Koebe function Moreover, it is easy to verify that the functions f maps $|z| < 1$ onto the w plane cut along the ray with constant argument from $-\frac{1}{4}e^{-i\alpha}$ to ∞. ●

The functions in (11.16) are extremal for Theorem 11.38 in the sense that there is equality on the bound for the second coefficient. Impressed by the fact that the Koebe function appears in many problems concerning the class $\mathcal{S}$, Bieberbach asked whether we always have $|a_n| \leq n$. This give rise to the famous

Bieberbach Conjecture. *If $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ is in $\mathcal{S}$, then $|a_n| \leq n$ for every n.*

Theorem 11.38 proves the conjecture for $n = 2$. Although stated in 1916, the conjecture was verified only for the values of n up to $n = 7$ until Louis de Branges proved the whole conjecture in 1985. For all n the maximization of $|a_n|$ is achieved only by the Koebe function. A large amount of research in the theory of univalent functions is centered on the Bieberbach conjecture.

The result for $n = 2$ can be used to prove the following elegant theorem which shows that this mapping property is, in a sense, extremal.

Theorem 11.40. *If $f(z) \in \mathcal{S}$ and $f(z) \neq c$ for $|z| < 1$, then $|c| \geq \frac{1}{4}$.*

Proof. Set $f(z) = z + a_2z^2 + \cdots$. Since $f(z) \neq c$, the function

$$g(z) = \frac{cf(z)}{c - f(z)} = z + \left(a_2 + \frac{1}{c}\right) z^2 + \cdots$$

is also in $\mathcal{S}$. Applying Theorem 11.38 to $g(z)$, we get $|a_2 + (1/c)| \leq 2$. Thus, $|1/c| - |a_2| \leq |(1/c) + a_2| \leq 2$. Now, applying Theorem 11.38 to $f(z)$, we have $|1/c| \leq 2 + |a_2| \leq 4$, and the result follows. ■

Remark 11.41. Theorem 11.40 is known as a *covering theorem* or *Koebe one-quarter theorem*. It says that every function in $\mathcal{S}$ maps the unit disk $|z| < 1$ onto a domain in the w plane that contains the disk $|w| < \frac{1}{4}$. This result has a lot of interesting applications in many other parts of complex analysis. By the inverse function theorem (also by the open mapping theorem), $f(|z| < 1)$ contains an open neighborhood of the origin (since $f(0) = 0$ and $f'(0) \neq 0$). The Koebe $\frac{1}{4}$–theorem actually estimates the size of this neighborhood. ●

Finally, we end the section with the following results which provides a sufficient condition for an analytic function to be univalent.

Theorem 11.42. *If $f(z)$ is analytic in a convex domain D, and* $\operatorname{Re} f'(z) > 0$ *in D, then $f(z)$ is univalent in D.*

Proof. Choose distinct points $z_0, z_1 \in D$. Then the straight line segment $z = z_0 + t(z_1 - z_0)$, $0 \le t \le 1$, must lie in D. Integrating along this path, we get

$$f(z_1) - f(z_0) = \int_{z_0}^{z_1} f'(z)\,dz = \int_0^1 f'(z_0 + t(z_1 - z_0))\,(z_1 - z_0)\,dt.$$

Dividing by $z_1 - z_0$ and taking real parts, we have

$$\operatorname{Re}\left\{\frac{f(z_1) - f(z_0)}{z_1 - z_0}\right\} = \operatorname{Re}\left\{\int_0^1 f'(z_0 + t(z_1 - z_0))\,dt\right\} > 0.$$

Thus $f(z_1) \ne f(z_0)$, and $f(z)$ is univalent in D. ∎

Questions 11.43.

1. What kind of results could have been obtained in this section if the functions had not been normalized?
2. What was the importance of the class $\mathcal{T}$?
3. Why was a bound on $|a_2|$ so useful?
4. Can $|a_2| = 2$ if $f(z)$ is a bounded function in $\mathcal{S}$?
5. Why is the Koebe function extremal for so many theorems?
6. For each n, are we guaranteed the existence of a function in $\mathcal{S}$ for which the absolute value of its nth coefficient is at least as large as the absolute value of the nth coefficient for any other function in $\mathcal{S}$?

Exercises 11.44.

1. Give an example of a function that is univalent but not analytic in the disk $|z| < 1$.
2. (a) If $f(z) \in \mathcal{S}$, show that for any nonzero complex number t, $|t| \le 1$, the function $f(tz)/t \in \mathcal{S}$.
   (b) If $f(z) = z/(1-z)^2$ and $|t_0| > 1$, show that $f(t_0 z)/t_0 \notin \mathcal{S}$.
3. If $f(z) = z + \sum_{k=2}^{\infty} a_k z^k$ is in $\mathcal{S}$, show that, for each integer n, there exists a function $g(z) = z + \sum_{k=2}^{\infty} b_k z^k$ in $\mathcal{S}$ such that $b_n = |a_n|$.
4. For α real, verify that $z/(1 - e^{i\alpha} z)^3$ is univalent in $|z| < \frac{1}{2}$, but in no larger disk centered at the origin.
5. If $f(z) \in \mathcal{S}$, show that $z(f(z^k)/z^k)^{1/k} \in \mathcal{S}$ for every positive integer k.
6. Let $f(z)$ be analytic in a domain D and suppose C is a closed contour in D. Prove that $\int_C \overline{f(z)} f'(z)\,dz$ is purely imaginary.
7. If $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ and $\sum_{n=2}^{\infty} n|a_n| \le 1$, show that $f(z) \in \mathcal{S}$.
8. If $f(z) = z - \sum_{n=2}^{\infty} |a_n| z^n$ is in $\mathcal{S}$, show that $\sum_{n=2}^{\infty} n|a_n| \le 1$.

12

Entire and Meromorphic Functions

We begin this chapter by defining infinite products, and show that their convergence properties are similar to those of infinite series. Just as infinite series were used as tools to develop power series expansions for analytic functions, so infinite products may be used as tools to develop product expansions for analytic functions. As we shall see, a comparison of product and series expansions enables us to determine some interesting identities.

Given any sequence $\{a_n\}$ tending to ∞, we show that there exists an entire function whose only zeros are at $\{a_n\}$. In one sense, this theorem is nicer than the Riemann mapping (existence) theorem of Section 11.3, because we can actually construct the entire function. We next consider an arbitrary sequence $\{b_n\}$ tending to ∞, and show that a function can always be found that has poles at $\{b_n\}$ and is analytic otherwise. In reading this chapter, it is worth keeping in mind the similarities between properties of zeros and poles.

12.1 Infinite Products

An infinite product is an expression of the form $u_1u_2u_3\cdots$ (denoted by $\prod_{n=1}^{\infty} u_n$), where the u_n are complex numbers. By analogy with infinite series, we are tempted to say that an infinite product converges if $\lim_{n\to\infty}(\prod_{k=1}^{n} u_k)$ exists. However, such a definition would be incomplete, because the vanishing of one term would necessitate the convergence of the infinite product regardless of the behavior of the other factors. This certainly is not in keeping with the spirit of "limit" definitions. We shall thus assume that no factor of an infinite product vanishes. Even so, we can (unlike the case for finite products) have an infinite product be zero although none of its factors is zero. For instance, $\lim_{n\to\infty}\prod_{k=1}^{n}(1/k) = 0$. Such an occurrence (the reasons for which will become evident later on) we also wish to avoid. We say that an infinite product $\prod_{n=1}^{\infty} u_n$ converges if and only if there is an N such that $u_k \neq 0$ for all $k \geq N$, $\lim_{n\to\infty}\prod_{k=N}^{n} u_k$ exists and is nonzero. An infinite product that does not converge is said to diverge. Moreover, if convergence condition holds but finitely many u_k's are equal to zero, then we say that the infinite product $\prod_{n=1}^{\infty} u_n$ converges to zero. In this case, the value of the product is set as 0. If $u_k \neq 0$ for all $k \geq 1$, and $\lim_{n\to\infty} \prod_{k=1}^{n} u_k = 0$, then we say that the infinite product diverges to zero.

To help clarify the definition of an infinite product, we make several simple observations. When discussing convergence, we need only to consider infinite products whose factors are all nonzero. Now, we set $P_n = \prod_{k=1}^{n} u_k$, where $u_k \neq 0$ for all $k \geq 1$. Here P_n is called the n-th partial product. If $\prod_{n=1}^{\infty} u_n$ converges, then the sequence $\{P_n\}$ approaches some nonzero value P. Thus,

$$u_n = \frac{P_n}{P_{n-1}} \to \frac{P}{P} = 1 \quad \text{as } n \to \infty.$$

It is therefore convenient to express a convergent infinite product in the form $\prod_{n=1}^{\infty}(1 + a_n)$, where $a_n \to 0$. We formulate

Theorem 12.1. (Necessary condition for convergence of a product) *If the infinite product $\prod_{n=1}^{\infty}(1+a_n)$ converges, then $a_n \to 0$ as $n \to \infty$, and $a_n = -1$ for at most finitely many n.*

As is the case with infinite series, the convergence of the sequence $\{a_n\}$ to 0 is not sufficient for the convergence of the product $\prod_{n=1}^{\infty}(1 + a_n)$. For example,

$$P_n = \prod_{k=1}^{n}\left(1 + \frac{1}{k}\right) = \frac{2}{1} \cdot \frac{3}{2} \cdot \frac{4}{3} \cdots \frac{n+1}{n} = n + 1 \to \infty,$$

so that $\prod_{n=1}^{\infty}(1 + 1/n)$ diverges; similarly,

$$P_n = \prod_{k=2}^{n+1}\left(1 - \frac{1}{k}\right) = \frac{1}{2} \cdot \frac{2}{3} \cdots \frac{n-1}{n} \cdot \frac{n}{n+1} = \frac{1}{n+1} \to 0,$$

so that $\prod_{n=2}^{\infty}(1 - 1/n)$ diverges.

It certainly seems natural to investigate further the comparison between the infinite series $\sum_{n=1}^{\infty} a_n$ and the infinite product $\prod_{n=1}^{\infty}(1 + a_n)$. In the special case that $a_n \geq 0$ for all n, the following relationship is particularly nice.

Theorem 12.2. *If $a_n \geq 0$, the product $\prod_{n=1}^{\infty}(1 + a_n)$ converges if and only if the series $\sum_{n=1}^{\infty} a_n$ converges.*

Proof. First note that

$$S_n := a_1 + a_2 + \cdots + a_n \leq P_n := (1 + a_1)(1 + a_2) \cdots (1 + a_n),$$

since all the terms are nonnegative. Also, $e^x \geq 1 + x$ for nonnegative x. Thus we have the double inequality

$$a_1 + a_2 + \cdots + a_n \le (1+a_1)(1+a_2)\cdots(1+a_n) \le e^{a_1}e^{a_2}\cdots e^{a_n}.$$

That is,

$$S_n \le P_n \le e^{S_n}. \tag{12.1}$$

If $\sum_{n=1}^{\infty} a_n$ converges to the real number S, then P_n is an increasing sequence that is bounded above (by e^S), and hence must converge. Conversely, if P_n converges to P, then the left-hand inequality of (12.1) shows that $\sum_{n=1}^{\infty} a_n$ converges to a value no greater than P. ■

For instance, if $a_n = O(1/n^p)$ as $n \to \infty$, then we know that the series $\sum_{n=1}^{\infty} a_n$ converges for $p > 1$ and diverges for $p \le 1$. Consequently, $\prod_{n=1}^{\infty}(1 + 1/n^p)$ converges for $p > 1$ and diverges for $p \le 1$.

As we shall see in the next theorem, similar results are obtained when the terms of the product are negative.

Theorem 12.3. *If $a_n \ge 0$, $a_n \ne 1$, then $\prod_{n=1}^{\infty}(1 - a_n)$ converges if and only if $\sum_{n=1}^{\infty} a_n$ converges.*

Proof. First suppose that $\sum_{n=1}^{\infty} a_n$ converges. By the Cauchy criterion, there exists N such that

$$a_N + a_{N+1} + \cdots + a_n < \frac{1}{2}$$

and $0 \le a_n < 1$ for all $n \ge N$. We have

$$(1 - a_N)(1 - a_{N+1}) = 1 - a_N - a_{N+1} + a_N a_{N+1} \ge 1 - a_N - a_{N+1} > \frac{1}{2}.$$

It can be shown by induction that for $n \ge N$,

$$\prod_{k=N}^{n}(1 - a_k) \ge 1 - \sum_{k=N}^{n} a_k > \frac{1}{2}. \tag{12.2}$$

Write

$$P_n = \prod_{k=1}^{n}(1 - a_k) = P_{N-1}\prod_{k=N}^{n}(1 - a_k).$$

Therefore, P_n/P_{N-1} is a decreasing sequence (since $0 < 1 - a_n \le 1$ for $n \ge N$) and has a lower bound. Thus, we get from (12.2) that P_n/P_{N-1} approaches a value P, $\frac{1}{2} \le P \le 1$. Thus $P_n \to P_{N-1}P \ne 0$, and hence $\prod_{k=1}^{\infty}(1 - a_k)$ converges.

To prove the converse, we suppose that $\sum_{n=1}^{\infty} a_n$ diverges. If $a_n \not\to 0$, then $1 - a_n \not\to 1$ and the product diverges. So we may assume, without loss of generality, that $a_n \to 0$. Then $0 \le a_n < 1$ for $n \ge N$. From the identity

$$e^{-x} = 1 - x + \left(\frac{x^2}{2!} - \frac{x^3}{3!}\right) + \left(\frac{x^4}{4!} - \frac{x^5}{5!}\right) + \cdots,$$

we see that $1 - x \le e^{-x}$ for $0 \le x < 1$, because all the terms in parenthesis are nonnegative. (Alternatively, it suffices to note that $f(x) = (1-x)e^x$ is decreasing on $[0, \infty)$ so that $f(x) \le f(0) = 1$). Hence

$$0 \le \prod_{k=N}^{n} (1 - a_k) \le \prod_{k=N}^{n} \exp(-a_k) = \exp\left(-\sum_{k=N}^{n} a_k\right) \quad (n \ge N).$$

Letting $n \to \infty$, the divergence of $\sum_{k=N}^{\infty} a_k$ shows that $\prod_{k=N}^{\infty}(1 - a_k) = 0$. Therefore $\prod_{k=1}^{\infty}(1 - a_k)$ diverges, and the proof is complete. ■

Remark 12.4. If a product were allowed to converge to 0, Theorem 12.3 would be false, as can be seen by letting $a_n = 1/(n+1)$. ●

When the restrictions on $\{a_n\}$ are relaxed, the comparison between $\prod_{n=1}^{\infty}(1+a_n)$ and $\sum_{n=1}^{\infty} a_n$ is less straightforward. In the exercises, we give an example for which $\prod_{n=1}^{\infty}(1+a_n)$ diverges even though $\sum_{n=1}^{\infty} a_n$ converges, and an example for which $\prod_{n=1}^{\infty}(1+a_n)$ converges even though $\sum_{n=1}^{\infty} a_n$ diverges.

In relating infinite products to infinite series in the general case, we will make use of the complex logarithm. If $\prod_{n=1}^{\infty}(1 + a_n) = P \ne 0$, then

$$\log\left[\prod_{n=1}^{\infty}(1 + a_n)\right] = \log P.$$

However, this does not mean that the series $\sum_{n=1}^{\infty} \log(1 + a_n)$ converges to $\log P$. Suppose

$$P_n = \prod_{k=1}^{n}(1 + a_k) = |P_n| e^{i \arg P_n}.$$

Then $|P_n| \to |P|$; but all we can say about $\arg P_n$ is that

$$\arg P_n \to \arg P \pmod{2\pi}.$$

In order to compare the convergence of $\prod_{n=1}^{\infty}(1 + a_n)$ with the convergence of $\sum_{n=1}^{\infty} \log(1 + a_n)$, it is necessary to deal with multiple-valued properties of the logarithm. The key step in the proof of our next theorem consists of showing that, for a convergent product, the arguments of the partial products eventually cluster about a fixed point when the same branch of $\arg(1 + a_k)$ is chosen for each k.

Theorem 12.5. *For a_n complex, $a_n \ne -1$, the product $\prod_{n=1}^{\infty}(1 + a_n)$ converges if and only if the series $\sum_{n=1}^{\infty} \operatorname{Log}(1 + a_n)$ converges.*

Proof. Set $P_n = \prod_{k=1}^{n}(1 + a_k)$, and write

$$\log P_n = \ln|P_n| + i \arg P_n = \sum_{k=1}^{n} \operatorname{Log}(1 + a_k) =: S_n, \tag{12.3}$$

where for each k branch of the logarithm has been chosen that satisfies $-\pi < \text{Im Log}\,(1 + a_k) = \text{Arg}\,(1 + a_k) \leq \pi$. This determines the branch of $\log P_n$.

Suppose that $P_n \to P \neq 0$. Then

$$\ln|P_n| \to \ln|P| \quad \text{and} \quad \arg P_n \to \arg P \ (\text{mod } 2\pi).$$

Thus there exists a sequence of real numbers $\{\epsilon_n\}$, $\epsilon_n \to 0$, and a sequence of integers $\{m_n\}$ such that for all n,

$$\arg P_n = \arg P + 2\pi m_n + \epsilon_n. \tag{12.4}$$

We will show that m_n is constant (say m) for n sufficiently large. To see this, by (12.3), we consider the difference

$$\log P_{n+1} - \log P_n = \text{ Log}\,(1 + a_{n+1})$$

so that

$$\arg P_{n+1} - \arg P_n = \text{Arg}\,(1 + a_{n+1}) = 2\pi(m_{n+1} - m_n) + \epsilon_{n+1} - \epsilon_n.$$

Since $a_n \to 0$, we have $\text{Arg}\,(1 + a_n) \to 0$. Hence for all $n > N$,

$$2\pi|m_{n+1} - m_n| \leq |\text{Arg}\,(1 + a_{n+1})| + |\epsilon_{n+1}| + |\epsilon_n| < 2\pi.$$

Therefore $m_{n+1} = m_n = m$ for $n > N$, and from (12.4) we see that

$$\arg P_n \to \arg P + 2m\pi.$$

In view of (12.3), it now follows that $\sum_{n=1}^{\infty} \text{Log}\,(1 + a_n)$ converges.

Conversely, suppose $S_n = \sum_{k=1}^{n} \text{Log}\,(1 + a_n)$. Then, we have

$$e^{S_n} = e^{\sum_{k=1}^{n} \text{Log}\,(1+a_k)} = e^{\text{Log}\,(1+a_1)} \cdots e^{\text{Log}\,(1+a_n)} = P_n. \tag{12.5}$$

Since the exponential function e^z is continuous, $S_n \to S$ implies that

$$e^{S_n} \to e^S \quad \text{as } n \to \infty.$$

Thus, letting $n \to \infty$ in (12.5) we find that $\prod_{k=1}^{\infty}(1 + a_k) = e^S \neq 0$. ■

Theorem 12.5 conveys that any question deal about infinite products can be translated into a question about infinite series by taking logarithms.

The infinite product $\prod_{n=1}^{\infty}(1 + a_n)$, $a_n \neq -1$, is said to be *absolutely convergent* if $\prod_{n=1}^{\infty}(1+|a_n|)$ converges. As in the case with series, an absolutely convergent product is convergent. Before proving this, we need the following.

Lemma 12.6. *For a_n complex, $a_n \neq -1$, $\sum_{n=1}^{\infty} a_n$ converges absolutely if and only if $\sum_{n=1}^{\infty}$ Log $(1 + a_n)$ converges absolutely. This occurs if and only if $\prod_{n=1}^{\infty}(1 + |a_n|)$ converges.*

Proof. If either of the two series converges, then there exists an N such that $|a_n| \le \frac{1}{2}$ for $n \ge N$. The Maclaurin expansion

$$\mathrm{Log}\,(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots = z\left(1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n+1} z^n\right)$$

is valid for $|z| < 1$. In particular, for $|z| \le \frac{1}{2}$,

$$\left|\sum_{n=1}^{\infty} \frac{(-1)^n}{n+1} z^n\right| \le \frac{1}{2}\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2}$$

which shows that

$$\frac{1}{2}|z| \le |\,\mathrm{Log}\,(1+z)| \le \frac{3}{2}|z| \quad \text{for } |z| \le 1/2.$$

Setting $z = a_n$, we have for $n \ge N$,

$$\frac{1}{2}|a_n| \le |\,\mathrm{Log}\,(1+a_n)| \le \frac{3}{2}|a_n|.$$

Hence either both series converge absolutely or neither series converges absolutely. ∎

It is easy to see that $\prod_{n=1}^{\infty}\left(1 + \frac{(-1)^{n+1}}{n}\right)$ converges to 1 (see Exercise 12.171(d) in 12.17) but not absolutely as $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.

Remark 12.7. In view of Lemma 12.6, Theorems 12.2 and 12.3 are seen to be special cases of Theorem 12.5. ●

Example 12.8. Consider the product $\prod_{k=1}^{\infty}(1 + 1/[k(2k+3)])$. Clearly (for example, by Theorem 12.5 and Lemma 12.6), the product converges. To find the limit value, we write the kth factor as

$$1 + \frac{1}{k(2k+3)} = \left(\frac{k+1}{k}\right)\left(\frac{2k+1}{2k+3}\right)$$

so that

$$P_n = 3\left(\frac{n+1}{2n+3}\right) \to \frac{3}{2}$$

as $n \to \infty$. Similarly, writing

$$1 + \frac{2}{k(k+3)} = \left(\frac{k+1}{k}\right)\left(\frac{k+2}{k+3}\right),$$

we see that the product $\prod_{k=1}^{\infty}(1 + 1/[k(k+3)])$ converges to 3 because

$$P_n = 3\left(\frac{n+1}{n+3}\right) \to 3 \quad \text{as } n \to \infty.$$

Finally, considering the equality

$$1-\frac{2}{(k+1)(k+2)}=\left(\frac{k}{k+1}\right)\left(\frac{k+3}{k+2}\right),$$

we obtain that the product $\prod_{k=1}^{\infty}(1-2/[(k+1)(k+2)])$ converges to $1/3$ because

$$P_n=\frac{1}{3}\left(\frac{n+3}{n+1}\right)\to\frac{1}{3}\text{ as }n\to\infty.$$

Note that the latter product is the inverse of the former. ●

Theorem 12.9. *If* $\prod_{n=1}^{\infty}(1+a_n)$ *converges absolutely, then* $\prod_{n=1}^{\infty}(1+a_n)$ *converges. The converse is not true.*

Proof. Suppose $\prod_{n=1}^{\infty}(1+|a_n|)$ converges. By Theorem 12.2, $\sum_{n=1}^{\infty}|a_n|$ converges. Then according to Lemma 12.6, $\sum_{n=1}^{\infty}|\operatorname{Log}(1+a_n)|$ converges. Since an absolutely convergent series is convergent, the result follows from Theorem 12.5.

The product $\prod_{n=1}^{\infty}(1+(-1)^{n+1}/n)$ converges to 1 but does not converge absolutely. ■

Remark 12.10. The terms of an absolutely convergent product can be rearranged without affecting the convergence or the value of the product. Its proof is similar to the comparable proof for infinite series, and is left as an exercise for the reader. ●

Example 12.11. Consider $a_k=(-1)^{k-1}/\sqrt{k+1}$. Then we see that $\sum_{k=1}^{\infty}a_k$ converges but $\prod_{k=1}^{\infty}(1+a_k)$ diverges, where $1+a_k\neq 0$ for each k. To show that the product diverges, we let

$$P_n=\prod_{k=1}^{n}\left(1+\frac{(-1)^{k-1}}{\sqrt{k+1}}\right).$$

Then

$$\begin{aligned}
P_{2n}&=\left(\frac{\sqrt{2}+1}{\sqrt{2}}\right)\left(\frac{\sqrt{3}-1}{\sqrt{3}}\right)\cdots\left(\frac{\sqrt{2n}+1}{\sqrt{2n}}\right)\left(\frac{\sqrt{2n+1}-1}{\sqrt{2n+1}}\right)\\
&\leq\left(\frac{\sqrt{3}+1}{\sqrt{2}}\,\frac{\sqrt{3}-1}{\sqrt{3}}\right)\cdots\left(\frac{\sqrt{2n+1}+1}{\sqrt{2n}}\,\frac{\sqrt{2n+1}-1}{\sqrt{2n+1}}\right)\\
&=\left(\frac{2}{\sqrt{2}\sqrt{3}}\right)\left(\frac{4}{\sqrt{4}\sqrt{5}}\right)\cdots\left(\frac{2n}{\sqrt{2n}\sqrt{2n+1}}\right)\\
&=\frac{1}{\sqrt{3/2}}\,\frac{1}{\sqrt{5/4}}\cdots\frac{1}{\sqrt{(2n+1)/2n}}\\
&=\frac{1}{\sqrt{\left(1+\frac{1}{2}\right)\left(1+\frac{1}{4}\right)\cdots\left(1+\frac{1}{2n}\right)}}.
\end{aligned}$$

As the product $\prod_{k=1}^{\infty}(1+1/(2k))$ diverges to ∞, it follows that $P_{2n} \to 0$ as $n \to \infty$. In particular, $\prod_{k=1}^{\infty}\left(1+\frac{(-1)^{k-1}}{\sqrt{k+1}}\right)$ is divergent, whereas the series $\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{\sqrt{k+1}}$, by alternating series test, converges but not absolutely. ●

Examples 12.12. We easily have

(i) $\displaystyle\prod_{k=1}^{\infty}\left(1+\frac{(-1)^k(1+i)}{k^3}\right)$ converges (since $\sum_{k=1}^{\infty} 1/k^3$ converges)

(ii) $\displaystyle\prod_{k=1}^{\infty}\left(1+\frac{i}{\sqrt{k(k+1)}}\right)$ diverges

(iii) $\displaystyle\prod_{k=1}^{\infty}\left(1+\frac{(-1)^k}{\sqrt{k}}\right)$ diverges

(iv) the product $\displaystyle\prod_{k=1}^{\infty}\left(1+\frac{i}{k}\right)$ and the series $\displaystyle\sum_{k=1}^{\infty}\operatorname{Log}(1+i/k)$ are divergent, whereas

$$\prod_{k=1}^{\infty}\left|1+\frac{i}{k}\right| = \prod_{k=1}^{\infty}\left(1+\frac{1}{k^2}\right)^{1/2}$$

converges, because the series $\sum_{k=1}^{\infty} k^{-2}$ converges. ●

Just as we went from series of complex numbers to series of functions, so may we go from products of complex numbers to products of complex functions. Also, as in the case of series of complex functions, the concept of uniform convergence plays an important role in the study of product of functions.

Let $\{f_n(z)\}_{n\geq 1}$ be a sequence of functions defined on a region Ω. Then the infinite product

$$\prod_{n=1}^{\infty}[1+f_n(z)]$$

is said to converge uniformly on Ω iff

(i) there exists an N such that $f_n(z) \neq -1$ for all $n > N$ and all $z \in \Omega$

(ii) the sequence $\prod_{k=N+1}^{n}[1+f_k(z)]$ converges uniformly on Ω to some $P(z)$, where $P(z) \neq 0$ for all $z \in \Omega$.

The most useful test for uniform convergence of products is analogous to the M-test (see Theorem 6.31) which has been used extensively to establish uniform (and absolute) convergence of series.

Theorem 12.13. **(M-test for the convergence of a product)** *Suppose that $\{f_n(z)\}$ is a sequence of functions such that $|f_n(z)| \leq M_n$ for all z in a region Ω. If $\sum_{n=1}^{\infty} M_n$ converges, then $\prod_{n=1}^{\infty}[1+f_n(z)]$ converges uniformly in Ω. In addition, if $f(z) = \prod_{n=1}^{\infty}[1+f_n(z)]$ and each $f_n(z)$ is analytic in Ω, then $f(z)$ is analytic in Ω. Also, $f(z_0) = 0$ for some $z_0 \in \Omega$ if and only if $f_n(z_0) = -1$*

for some n. The order of the zero of f at z_0 is the sum of the order of zeros of the functions $1 + f_n(z)$ at z_0.

Proof. According to Theorem 12.2, $\prod_{n=1}^{\infty}[1 + f_n(z)]$ converges absolutely, hence converges, for each point in Ω. It suffices to show that the sequence $P_n(z) = \prod_{k=1}^{n}[1 + f_k(z)]$ is uniformly Cauchy in Ω. Note that for any positive integers m and n ($m < n$), we have

$$P_n(z) - P_m(z) = \sum_{k=m}^{n-1}[P_{k+1}(z) - P_k(z)] = \sum_{k=m}^{n-1} P_k(z) f_{k+1}(z). \tag{12.6}$$

But for all k,

$$|P_k(z)| \leq \prod_{n=1}^{\infty}(1 + |f_n(z)|) \leq \exp \sum_{n=1}^{\infty} |f_n(z)| \leq \exp \sum_{n=1}^{\infty} M_n := e^M.$$

By Theorem 6.26, the series $\sum_{n=1}^{\infty} |f_n(z)|$ converges uniformly in R. Hence, choosing m large enough in (12.6) so that $\sum_{k=m}^{n-1} |f_{k+1}(z)| < \epsilon$ for all $z \in \Omega$ and for all n, we have

$$|P_n(z) - P_m(z)| \leq \sum_{k=m}^{n-1} |P_k(z)|\,|f_{k+1}(z)| \leq e^M \sum_{k=m}^{n-1} |f_{k+1}(z)| < \epsilon e^M.$$

Since ϵ is arbitrary, it follows that the sequence $\{P_n(z)\}$ is uniformly Cauchy in Ω and the proof of the first part is complete.

Next suppose that $f(z_0) = 0$ for some $z_0 \in \Omega$. Then by the definition of the convergence of infinite products, there exists an N such that

$$F_N(z) = \prod_{k=N+1}^{\infty} (1 + f_k(z))$$

is nonvanishing at z_0. By the above discussion, $F_N(z)$ is the limit of a uniformly convergent sequence of analytic functions. Hence the limit $F_N(z)$ is analytic in Ω. Continuity of $F_N(z)$ at z_0 shows that $F_N(z)$ is nonvanishing in some neighborhood $D(z_0; \delta)$ of z_0. Now

$$f(z) = \prod_{k=1}^{\infty}(1 + f_k(z)) = \prod_{k=1}^{N}(1 + f_k(z)) \left(\lim_{n\to\infty} \prod_{k=N+1}^{n} (1 + f_k(z)) \right).$$

Note that the second factor is analytic and nonvanishing on $D(z_0; \delta)$. Therefore, the zero of $f(z)$ and their order arise only from the zeros of the factor $\prod_{k=1}^{N}(1 + f_k(z))$. ∎

Remark 12.14. Showing the sequence $\{P_n(z)\}$ to be uniformly Cauchy does not preclude the possibility that $P_n(z) \to 0$ for some z. This is why it was necessary to show that $\{P_n(z)\}$ also converged pointwise (to a nonvanishing function).

Example 12.15. Consider the product $\prod_{k=0}^{\infty}(1+z^{2^k})$. Clearly, the series $\sum_{k=0}^{\infty} z^{2^k}$ converges absolutely for $|z|<1$, and hence the product converges absolutely for $|z|<1$. Now we observe that

$$\begin{aligned}
(1-z)P_0(z) &= 1-z^2 \\
(1-z)P_1(z) &= (1-z^2)(1+z^2) = 1-z^{2^2} \\
&\vdots \qquad\qquad \vdots \\
(1-z)P_n(z) &= (1-z^{2^n})(1+z^{2^n}) = 1-z^{2^{n+1}}
\end{aligned}$$

and hence, for $|z|<1$, we have

$$(1-z)\lim_{n\to\infty} P_n(z) = \lim_{n\to\infty}(1-z^{2^{n+1}}) = 1,$$

as desired. More generally, we have

$$\prod_{k=0}^{\infty}\left(1+\left(\frac{z}{R}\right)^{2^k}\right) = \frac{R}{R-z} \quad \text{for } |z|<R.$$

In particular,

$$\prod_{k=1}^{\infty}\left(1+\left(\frac{z}{R}\right)^{2^k}\right) = \frac{R^2}{R^2-z^2} \quad \text{for } |z|<R.$$

●

Questions 12.16.

1. What are the similarities between infinite series and infinite products?
2. If a sequence of partial products $\{P_n\}$ converges, does $\{\log P_n\}$ converge? What about the converse?
3. In view of Theorem 12.2 and Theorem 12.3, is it true for real sequences $\{a_n\}$ that $\prod_{n=1}^{\infty}(1+a_n)$ converges if and only if $\sum_{n=1}^{\infty} a_n$ converges?
4. How many ways can an infinite product diverge?
5. If $\prod_{n=1}^{\infty}(1+a_n)$ and $\prod_{n=1}^{\infty}(1+b_n)$ converge, does $\prod_{n=1}^{\infty}(1+a_n)(1+b_n)$ also converge?
6. If $\prod_{n=1}^{\infty}(1+a_n)$ and $\prod_{n=1}^{\infty}(1+b_n)$ diverge, does $\prod_{n=1}^{\infty}(1+a_n)(1+b_n)$ diverge? Can $\prod_{n=1}^{\infty}(1+a_n)(1+b_n)$ be convergent?
7. If α and β are two real numbers such that $\alpha+\beta=-1$, does the product $\prod_{n=1}^{\infty}\left(1+n^{\alpha}(1+n)^{\beta}\right)$ converge?
8. We know that a series converges if and only if every subsequence obtained by deleting a finite number of terms of the series converges. Does the definition of infinite product assure the analogous statement for infinite products?
9. If $\prod_{n=1}^{\infty}(1+a_n)$ converges, does $\prod_{n=1}^{\infty}|1+a_n|$ converge?
10. If $\prod_{n=1}^{\infty}(1+a_n)$ and $\prod_{n=1}^{\infty}(1+|a_n|)$ diverge, what can we say about the sequence $\{a_n\}$?

11. Can $\prod_{n=1}^{\infty}(1+f_n(z))$ converge uniformly but not absolutely? Absolutely but not uniformly?
12. What would we mean by an infinite product satisfying the Cauchy criterion? What is its relation to convergence?
13. If $\sum_{n=1}^{\infty}|a_n|$ converges, what can be said about $\prod_{n=1}^{\infty}(1+a_n z)$? How about $\prod_{n=1}^{\infty}(1+a_n z^2)$? How about $\prod_{n=1}^{\infty}(1+a_n p(z))$, where $p(z)$ is some polynomial?
14. How would Theorem 12.5 be proved without making the assumption that $-\pi < \arg(1+a_k) \leq \pi$ for each k?

Exercises 12.17.

1. Show that
   (a) $\prod_{n=2}^{\infty}\left(1-\dfrac{1}{n^2}\right)=\dfrac{1}{2}$ (b) $\prod_{n=1}^{\infty}\left(1+\dfrac{1}{n(n+2)}\right)=2$
   (c) $\prod_{n=2}^{\infty}\left(1-\dfrac{2}{n(n+1)}\right)=\dfrac{1}{3}$ (d) $\prod_{n=1}^{\infty}\left(1+\dfrac{(-1)^{n+1}}{n}\right)=1$
   (e) $\prod_{n=3}^{\infty}\dfrac{n^2-1}{n^3-4}=4$ (f) $\prod_{n=2}^{\infty}\dfrac{n^3-1}{n^3+1}=\dfrac{2}{3}$.
2. Show that the product
$$\left(1-\frac{1}{2}\right)\left(1+\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1+\frac{1}{3}\right)\left(1-\frac{1}{4}\right)\cdots$$
   converges, but not absolutely.
3. Suppose $\{a_n\}$ is real with $|a_n|<1$. If $\sum_{n=1}^{\infty}a_n$ converges, show that $\prod_{n=1}^{\infty}(1+a_n)$ converges if and only if $\sum_{n=1}^{\infty}a_n^2$ converges.
4. Set
$$a_n=\begin{cases}\dfrac{1}{\sqrt{n}}+\dfrac{1}{n}+\dfrac{1}{n\sqrt{n}} & \text{if } n \text{ even},\\[2ex] -\dfrac{1}{\sqrt{n}} & \text{if } n \text{ odd}.\end{cases}$$
   Show that $\prod_{n=1}^{\infty}(1+a_n)$ converges but both $\sum_{n=1}^{\infty}a_n$ and $\sum_{n=1}^{\infty}a_n^2$ diverge.
5. Suppose that $\{a_n\}$ is a decreasing sequence of real numbers, with
$$\lim_{n\to\infty}a_n=0.$$
   Show that $\prod_{n=1}^{\infty}[1+(-1)^n a_n]$ converges if and only if $\sum_{n=1}^{\infty}a_n^2$ converges.
6. Set
$$a_n=\begin{cases}\dfrac{1}{\sqrt{n}}+\dfrac{1}{n} & \text{if } n \text{ odd},\\[2ex] -\dfrac{1}{\sqrt{n}} & \text{if } n \text{ even}.\end{cases}$$

Show that $\prod_{n=1}^{\infty}(1+a_n)$ converges even though $\sum_{n=1}^{\infty} a_n$ diverges. Does this provide an example of an infinite product that is convergent but not absolutely?

7. Show that the following either all converge or all diverge:

$$\sum_{n=1}^{\infty} |a_n|, \quad \prod_{n=1}^{\infty}(1+|a_n|), \quad \sum_{n=1}^{\infty} |\log(1+a_n)|, \quad \sum_{n=1}^{\infty} \log(1+|a_n|).$$

8. Determine the region of convergence for
   (a) $\prod_{n=1}^{\infty}(1+z^n)$ (b) $\prod_{n=1}^{\infty}(1-z^{2^n})$ (c) $\prod_{n=2}^{\infty}(1-n^{-z})$
   (d) $\prod_{n=1}^{\infty}\cos(z/n)$ (e) $\prod_{n=1}^{\infty}\sin(z/n)$ (f) $\prod_{n=2}^{\infty}\cos(z/2^n)$.
9. Suppose $a_n > 0$ for every n, and $\sum_{n=1}^{\infty} a_n$ converges. Show that

$$a_1 \prod_{n=2}^{\infty}\left(1+\frac{a_n}{s_{n-1}}\right) = \sum_{n=1}^{\infty} a_n,$$

where $s_n = \sum_{k=1}^{n} a_k$.

12.2 Weierstrass' Product Theorem

Let us now consider the factorization of entire functions. As a first step consider an entire function f which does not vanish in $\mathbb{C}$. Then we may express $f(z)$ as

$$f(z) = \exp(g(z)),$$

where $g(z)$ is an entire function. In fact, as $f'(z)/f(z)$ is analytic in $\mathbb{C}$, $f'(z)/f(z)$ possesses an anti-derivative $g(z)$ so that

$$g'(z) = \frac{f'(z)}{f(z)}$$

for some entire function g. Using this, we see that

$$(f(z)e^{-g(z)})' = 0, \quad \text{i.e.,} \quad f(z) = c\exp(g(z))$$

for some constant c. Absorbing the constant c into $g(z)$, we obtain the desired representation.

Next, given a finite set of complex numbers $\{0, a_1, a_2, \dots, a_n\}$ ($a_k \neq 0$ for $1 \leq k \leq n$), we can always find an entire function (in fact, a polynomial) having zeros at these points of order $m, m_1, \dots, m_n$, respectively. Since such an entire function is analytic in $\mathbb{C}$ with no singularities except at ∞, one such entire function is the polynomial

$$p(z) = z^m \prod_{k=1}^{n}\left(1-\frac{z}{a_k}\right)^{m_k}.$$

Moreover, if f is an entire function with a finite number of zeros at $0, a_k$ ($a_k \neq 0$) for $1 \leq k \leq n$, of order m, m_k ($1 \leq k \leq n$), respectively, then

$$h(z) = \frac{f(z)}{p(z)}$$

has removable singularities at $0, a_k$ ($1 \leq k \leq n$). It follows then that $h(z)$ defines an entire function with no zeros in $\mathbb{C}$. As a consequence,

$$f(z) = e^{g(z)} p(z) = e^{g(z)} z^m \prod_{k=1}^{n} \left(1 - \frac{z}{a_k}\right)^{m_k}.$$

That is, $f(z)$ can be expressed as a product of a polynomial and a nonvanishing entire function.

Finally, the question arises as to whether an entire function can always be found whose only zeros are at an arbitrarily prescribed sequence of points. Unfortunately, the answer is no. For example, if an entire function has zeros at $1/n$ ($n \in \mathbb{N}$), then according to Theorem 8.47, the entire function must be identically zero. More generally, a nonconstant entire function cannot have a limit point of zeros in $\mathbb{C}$. Consequently, the set of zeros of an entire function which has infinitely many zeros in $\mathbb{C}$ must have ∞ as its only limit point. For example,

$$\begin{aligned}
0 = \cos z = \frac{e^{iz} + e^{-iz}}{2} &\implies e^{2iz} = -1 = e^{i(\pi + 2k\pi)} \\
&\implies z = (2k+1)\pi/2, \quad k \in \mathbb{Z}; \\
0 = \sin z = \frac{e^{iz} - e^{-iz}}{2} &\implies e^{2iz} = 1 = e^{2k\pi i}, \text{ i.e., } z = k\pi, \quad k \in \mathbb{Z},
\end{aligned}$$

and $e^z - 1 = 0 \implies z = 2k\pi i$ for $k \in \mathbb{Z}$. Thus, in each case, the limit point of the zeros of $\cos z$, $\sin z$, and e^z is ∞. We will show, however, that if the sequence of zeros of an entire function has no finite limit point, then the question concerning factorization can be answered in the affirmative.

Suppose a sequence $\{a_n\}_{n \geq 1}$ approaching ∞ is arranged so that

$$0 < |a_1| \leq |a_2| \leq |a_3| \leq \cdots.$$

A naive guess for an appropriate entire function is

$$f(z) = \prod_{n=1}^{\infty} \left(1 - \frac{z}{a_n}\right).$$

Unfortunately, $\prod_{n=1}^{\infty}(1 - z/a_n)$ may diverge. Certainly if $f(z)$ is entire, then the function has its only zeros at points of the sequence $\{a_n\}$ and nowhere else.

Example 12.18. Let us first discuss an example with some details. Consider the infinite product $\prod_{n=1}^{\infty}(1-z^2/n^2)$ which has zeros at $z=\pm n$, $n\in\mathbb{N}$. Fix an arbitrary $R>0$. If $|z|\leq R$, then choose N large enough so that

$$|z/n|\leq R/n<1 \quad \text{for all } n\geq N,$$

and so, $1-z^2/n^2\neq 0$ for $|z|\leq R$ and for $n\geq N$. In view of this observation, we may write

$$\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2}\right)=P_{N-1}(z)\prod_{n=N}^{\infty}\left(1-\frac{z^2}{n^2}\right)=P_{N-1}(z)F_N(z).$$

Note that $P_{N-1}(z)$ is entire and has zeros only at the points $z=\pm n$ $(n<N)$ and $F_N(z)$ is an infinite product with no zeros in $|z|\leq R$. Also,

$$\sum_{n=1}^{\infty}\left|\frac{z^2}{n^2}\right|=|z|^2\sum_{n=1}^{\infty}\frac{1}{n^2}\leq R^2\sum_{n=1}^{\infty}\frac{1}{n^2}<\infty$$

and so, by Theorem 12.13, the infinite product is uniformly convergent for $|z|\leq R$ and hence, on every compact subsets of $\mathbb{C}$. By Theorem 12.13, we conclude that $\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2}\right)$ is entire, and the infinite product is zero only at $z=\pm n$, $n\in\mathbb{N}$. ●

We shall now determine the restrictions on $\{a_n\}$ for which $f(z)$ is entire. Suppose that $\sum_{n=1}^{\infty}1/|a_n|$ converges. Fix an arbitrary $R>0$. If $|z|\leq R$, then $|z/a_n|\leq R/|a_n|$ and an application of Theorem 12.13 shows that $P_n(z)$ converges uniformly to $f(z)=\prod_{n=1}^{\infty}(1-z/a_n)$ in $|z|\leq R$ and hence, on every compact subset of $\mathbb{C}$. By Theorem 8.16, $f(z)$ must therefore be an entire function. For example, we can now easily construct an entire function whose only zeros are at $1,4,9,16,\ldots$. The product $\prod_{n=1}^{\infty}(1-z/n^2)$ is such a function.

However, it is more difficult to construct an entire function whose zeros are at the positive integers. The expression $\prod_{n=1}^{\infty}(1-z/n)$ does not represent an entire function. In fact, setting $z=-1$ we see that

$$\lim_{n\to\infty}\prod_{k=1}^{n}\left(1+\frac{1}{k}\right)=\lim_{n\to\infty}(n+1)=\infty.$$

What is needed is a "convergence producing" factor. Moreover, if the series $\sum_{n=1}^{\infty}1/|a_n|$ diverges but $\sum_{n=1}^{\infty}1/|a_n|^2$ converges, we can modify the above construction and show that

$$\prod_{n=1}^{\infty}\left(1-\frac{z}{a_n}\right)e^{z/a_n}$$

is entire. We will first show that

$$f(z) = \prod_{n=1}^{\infty} \left(1 - \frac{z}{n}\right) e^{z/n}$$

is an entire function and the same method can be adapted to solve a more general problem. We set

$$1 + f_n(z) = (1 - z/n)e^{z/n}.$$

To determine the uniform convergence of the product, we need to find an upper bound for $|f_n(z)|$. We write

$$1 + f_n(z) = \exp\left(\operatorname{Log}(1 - z/n) + z/n\right).$$

Fix an arbitrary $R > 0$. If $|z| \leq R$, then choose N large enough so that $N \geq 2R$. Then $|z/n| \leq R/n < 1$ for all $n \geq N$, and so the identity

$$\operatorname{Log}\left(1 - \frac{z}{n}\right) = -\sum_{k=1}^{\infty} \frac{1}{k}\left(\frac{z}{n}\right)^k \tag{12.7}$$

is valid for $|z| \leq R$. We get

$$\begin{aligned}\left|\operatorname{Log}\left(1 - \frac{z}{n}\right) + \frac{z}{n}\right| &= \left|-\sum_{k=2}^{\infty} \frac{1}{k}\left(\frac{z}{n}\right)^k\right| \leq \frac{1}{2}\sum_{k=2}^{\infty}\left(\frac{R}{n}\right)^k \\ &= \frac{1}{2}\frac{R^2/n^2}{1 - R/n} \leq \frac{R^2}{n^2},\end{aligned}$$

because $R < 2R \leq N \leq n$. Hence,

$$\begin{aligned}|f_n(z)| &= |\exp\left(\operatorname{Log}(1 - z/n) + z/n\right) - 1| \\ &\leq \exp\left(|\operatorname{Log}(1 - z/n) + z/n|\right) - 1 \\ &\leq \exp(R^2/n^2) - 1 \\ &\leq (R^2/n^2)\exp(R^2/n^2) \quad (\text{since } e^x - 1 \leq xe^x \text{ for } x \geq 0) \\ &\leq e(R^2/n^2) = M_n \quad (\text{since } R/n < 1)\end{aligned}$$

so that

$$\sum_{n=N}^{\infty} |f_n(z)| \leq eR^2 \sum_{n=N}^{\infty} \frac{1}{n^2} < \infty \quad (N > 2R).$$

According to Theorem 12.5, $\prod_{n=1}^{\infty}(1 + f_n(z))$ converges uniformly for $|z| \leq R$. Hence

$$f(z) = \prod_{n=1}^{\infty}\left(1 - \frac{z}{n}\right)e^{z/n} = \prod_{n=1}^{N-1}\left(1 - \frac{z}{n}\right)e^{z/n} \prod_{n=N}^{\infty}\left(1 - \frac{z}{n}\right)e^{z/n}$$

is an entire function with the prescribed zeros. Another such function may be obtained just by multiplying $f(z)$ by any nonvanishing entire function.

Similarly, $f(z) = \prod_{n=1}^{\infty}(1 + z/n)e^{-z/n}$ is an entire function with simple zeros at the negative integers and no other zeros. Consequently, an entire function having a simple zero at each integer is given by

$$\begin{aligned} z\prod_{n=1}^{\infty}\left(1-\frac{z}{n}\right)e^{z/n}\prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)e^{-z/n} &= z\prod_{n=1}^{\infty}\left(1-\frac{z}{n}\right)\left(1+\frac{z}{n}\right)e^{z/n-z/n} \\ &= z\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2}\right). \end{aligned}$$

The rearrangement of the factors in the first equality is justified by the absolute convergence of the infinite products.

A more general method for constructing an entire function with prescribed zeros is indicated by the identity (12.7). For instance, suppose we want an entire function to have its zeros at $z = \sqrt{n}$, $n \in \mathbb{N}$. Since

$$\mathrm{Log}\left(1-\frac{z}{\sqrt{n}}\right) = -\frac{z}{\sqrt{n}} - \frac{1}{2}\left(\frac{z}{\sqrt{n}}\right)^2 - \frac{1}{3}\left(\frac{z}{\sqrt{n}}\right)^3 - \cdots,$$

the above method shows that

$$\prod_{n=1}^{\infty}\left(1-\frac{z}{\sqrt{n}}\right)e^{(z/\sqrt{n})+(1/2)(z^2/n)}$$

is such a function. Similarly, an entire function that has simple zeros on the real axis at points $z = \pm n^{1/4}$ $(n \geq 0)$ and nowhere else is given by

$$z\prod_{n=1}^{\infty}\left(1-\frac{z^2}{\sqrt{n}}\right)e^{z^2/\sqrt{n}+(1/2)(z^4/n)}.$$

More generally, if $\sum_{n=1}^{\infty} 1/|a_n|^{p+1}$ converges for some positive integer p, then

$$f(z) = \prod_{n=1}^{\infty}\left(1-\frac{z}{a_n}\right)E_p\left(\frac{z}{a_n}\right) \tag{12.8}$$

is an entire function whose only zeros are at $z = a_n$, where

$$E_p(z) = \exp\left(z + (1/2)z^2 + \cdots + (1/p)z^p\right)$$

and is referred to as the convergence producing factor.

As general as (12.8) appears, we still have not accounted for all sequences. For instance, suppose we wish to construct an entire function whose zeros occur at the points $\log n$ $(n = 2, 3, 4, \ldots)$. We cannot use (12.8) because $\sum_{n=2}^{\infty}[1/(\ln n)^p]$ diverges for all p. Observe that the convergence producing factors in (12.8) involve a sequence of polynomials, all of degree p. In the general case, we will not place a uniform bound on the degree of the polynomials. This, in turn, will enable us to construct an appropriate entire function without regard to the convergence of a series involving its zeros.

Theorem 12.19. (Weierstrass's Product Theorem) *Given any complex sequence having no finite limit point, there exists an entire function that has zeros at these points and only these points.*

Proof. We suppose that the entire function $f(z)$ to be constructed is to have zeros at $\{a_n\}$ so arranged that $0 < |a_1| \leq |a_2| \leq |a_3| \leq \cdots$. We have assumed, without loss of generality, that none of the a_n is zero, for if the k of these points are zero, then $z^k f(z)$ has the desired property.

For each n, set

$$P_n(z) = \frac{z}{a_n} + \frac{1}{2}\left(\frac{z}{a_n}\right)^2 + \cdots + \frac{1}{n}\left(\frac{z}{a_n}\right)^n$$

so that $\exp(P_n(z)) = E_n(z/a_n)$, where $E_n(z)$ is the convergence producing factor as defined above. We wish to show that the function

$$f(z) = \prod_{n=1}^{\infty}\left(1 - \frac{z}{a_n}\right) E_n\left(\frac{z}{a_n}\right)$$

satisfies the conditions of the theorem. As in the proof of above example, it suffices to show that

$$\sum_{n=1}^{\infty}\left[\mathrm{Log}\left(1 - \frac{z}{a_n}\right) + P_n(z)\right]$$

converges uniformly on an arbitrary compact subset $|z| \leq R$ of the plane. Choose $|a_n|$ large enough so that $|a_n| \geq 2R \geq 2|z|$. Then, we have

$$\begin{aligned}\left|\mathrm{Log}\left(1 - \frac{z}{a_n}\right) + P_n(z)\right| &= \left|-\sum_{k=n+1}^{\infty}\frac{1}{k}\left(\frac{z}{a_n}\right)^k\right| \\ &\leq \frac{1}{n+1}\sum_{k=n+1}^{\infty}\left|\frac{z}{a_n}\right|^k \\ &\leq \frac{1}{n+1}\left|\frac{z}{a_n}\right|^n \sum_{k=1}^{\infty}\frac{1}{2^k} \\ &\leq \left|\frac{z}{a_n}\right|^n \leq \frac{1}{2^n}.\end{aligned}$$

If we set $1 + f_n(z) = \exp\left(\mathrm{Log}\,(1 - z/a_n) + P_n(z)\right)$, then we have

$$|f_n(z)| \leq \exp(1/2^n) - 1 \leq (1/2^n)e.$$

According to Theorem 12.5, $\prod_{n=1}^{\infty}(1 + f_n(z))$ converges uniformly to $f(z)$ for $|z| \leq R$. Since R was arbitrary, the infinite product defines an entire function with prescribed zeros. The assertion in the theorem about the zeros of $f(z)$ follows from the definition. ∎

Remark 12.20. The theorem does not exclude the case of multiple zeros. It is certainly possible that $a_k = a_{k+1}$ for some k. ●

Remark 12.21. The function $f(z)$ is by no means uniquely determined by the zeros. For instance, we have seen that $\prod_{n=1}^{\infty}(1 - z/n)e^{z/n}$ is an entire function having zeros at the positive integers. In the proof of the theorem, we have shown that

$$\prod_{n=1}^{\infty}\left(1 - \frac{z}{n}\right) e^{(z/n)+(1/2)(z/n)^2+\cdots+(1/n)(z/n)^n}$$

is another such function. We can also show that

$$\prod_{n=1}^{\infty}\left(1 - \frac{z}{n}\right) e^{(z/n)+(1/2)(z/n)^2}, \quad \prod_{n=1}^{\infty}\left(1 - \frac{z}{n}\right) e^{(z/n)+(1/2)(z/n)^2+(1/3)(z/n)^3},$$

and infinitely many more entire functions also have their only zeros at the positive integers. ●

Suppose two entire functions have the same zeros with the same multiplicities. How do the two functions compare? Since the complex plane $\mathbb{C}$ is simply connected, our characterization is a consequence of Theorem 7.51. We have

Theorem 12.22. *If $f(z)$ and $g(z)$ are entire functions whose zeros coincide in location and in multiplicity, then there exists an entire function $\phi(z)$ such that $f(z) = e^{\phi(z)}g(z)$.*

Proof. After we cancel the common factors, the function $f(z)/g(z)$ is seen to be an entire function with no zeros. The result follows from Theorem 7.51 (see also the discussion in the beginning). ■

Weierstrass's theorem shows that for a preassigned sequence of points, we can construct an infinite product that represents an entire function having zeros at the preassigned sequence. What about the converse? Given an entire function whose zeros are known, can we construct an infinite product representation for the function? In view of Theorem 12.22, this is always possible up to a multiplicative exponential function. The explicit determination of the exponential function is usually quite difficult. Before we solve this problem for the function $\sin \pi z$, it is appropriate to formulate Theorem 12.22 in the following form.

Theorem 12.23. *Let $\{a_n\}_{n\geq 1}$ be a sequence of nonzero complex numbers and $f(z)$ be an entire function that has zeros at a_n, listed with multiplicities. Suppose that f has a zero of order $k \geq 0$ at zero. Then there exists an entire function $g(z)$ such that*

$$f(z) = z^k e^{g(z)} \prod_{n=1}^{\infty}\left(1 - \frac{z}{a_n}\right) E_n\left(\frac{z}{a_n}\right).$$

More generally, one can replace the product in the last equation by

$$\prod_{n=1}^{\infty}\left(1-\frac{z}{a_n}\right)E_{p_n}\left(\frac{z}{a_n}\right)$$

where $\{p_n\}_{n\geq 1}$ *is a sequence in* $\mathbb{N}$ *such that for each* $R>0$,

$$\sum_{n=1}^{\infty}\left(\frac{R}{|a_n|}\right)^{p_n+1}<\infty.$$

Suppose that $P_n(z)=\prod_{k=1}^{n}[1+f_k(z)]$ is a finite product of analytic functions on a domain D. Then logarithmic differentiation gives

$$\frac{P_n'(z)}{P_n(z)}=\sum_{k=1}^{n}\frac{f_k'(z)}{1+f_k(z)}.$$

Note that $P_n'(z)/P_n(z)$ has poles at the zeros of $P_n(z)$. This procedure continues to hold for uniformly convergent infinite products of analytic functions and we state the following result whose proof is left as a simple exercise.

Theorem 12.24. *Suppose* $\{f_n(z)\}_{n\geq 1}$ *is a sequence of analytic functions in a domain* D *and let* $\prod_{n=1}^{\infty}[1+f_n(z)]$ *converge uniformly on* D *to* $f(z)$. *Then*

$$\frac{f'(z)}{f(z)}=\sum_{k=1}^{\infty}\frac{f_k'(z)}{1+f_k(z)},$$

where the sum converges uniformly on D *when* $f(z)\neq 0$.

The following result is often useful for expanding entire or meromorphic functions.

Theorem 12.25. *Let* f *be analytic except for simple poles at* $a_1,a_2,\ldots$ *and be arranged so that*

$$0<|a_1|\leq|a_2|\leq\cdots\leq|a_n|\leq\cdots,\qquad \text{with } b_n=\operatorname{Res}[f(z);a_n].$$

Let $\{C_n\}$ *be a sequence of positively oriented simple closed contours such that each* C_n *includes* $a_1,a_2,\ldots,a_n$ *but no other poles. Suppose that*

$$\begin{aligned} R_n &= \operatorname{dist}(0,C_n)\to\infty \quad \text{as } n\to\infty,\\ L_n &= \text{length of } C_n = O(R_n),\\ |f(z)| &= o(R_n) \quad \text{on } C_n \end{aligned}$$

(*e.g., the last condition is satisfied if* $f(z)$ *is bounded on all* C_n). *Then,*

$$f(z)=f(0)+\sum_{n=1}^{\infty}b_n\left(\frac{1}{z-a_n}+\frac{1}{a_n}\right)$$

for all z *except at these poles.*

Proof. Define

$$F_n(\alpha) = \frac{1}{2\pi i}\int_{C_n} g(z)\,dz, \quad g(z) = \frac{f(z)}{z(z-\alpha)},$$

where α lies inside C_n. If α is not a pole of f, then $g(z)$ has simple poles at each $a_k, 0, \alpha$ with

$$\begin{aligned}
\operatorname{Res}[g(z); a_k] &= \lim_{z\to a_k}(z-a_k)\frac{f(z)}{z(z-\alpha)} = \frac{b_k}{a_k(a_k-\alpha)}\\
\operatorname{Res}[g(z); 0] &= \lim_{z\to 0} z\frac{f(z)}{z(z-\alpha)} = -\frac{f(0)}{\alpha}\\
\operatorname{Res}[g(z); \alpha] &= \lim_{z\to \alpha}(z-\alpha)\frac{f(z)}{z(z-\alpha)} = \frac{f(\alpha)}{\alpha},
\end{aligned}$$

respectively. Therefore, by the Residue Theorem,

$$F_n(\alpha) = \sum \operatorname{Res}[g(z); C_n]$$

which gives

$$F_n(\alpha) = \sum_{k=1}^{n} \frac{b_k}{a_k(a_k-\alpha)} - \frac{f(0)}{\alpha} + \frac{f(\alpha)}{\alpha}. \tag{12.9}$$

Now, for $z \in C_n$,

$$|z| \geq R_n = \operatorname{dist}(0, C_n) \quad \text{and} \quad |z-\alpha| \geq |z| - |\alpha| \geq R_n - |\alpha| > 0$$

so that

$$|F_n(\alpha)| \leq \frac{1}{2\pi}\frac{L_n}{R_n(R_n-|\alpha|)}\max_{z\in C_n}|f(z)| \to 0 \quad \text{as } n\to\infty\ ,$$

and therefore the sequence $\{F_n(\alpha)\}$ converges to zero uniformly on the compact set C_n. Allowing $n \to \infty$ in (12.9) gives

$$f(\alpha) = f(0) + \sum_{k=1}^{\infty} b_k\left(\frac{1}{\alpha - a_k} + \frac{1}{a_k}\right). \qquad \blacksquare$$

Now we recall that $\sin \pi z$ has simple zeros at all the integers. We have seen that

$$z\prod_{n=1}^{\infty}\left(1 - \frac{z^2}{n^2}\right)$$

is an entire function having a simple zero at each integer. Therefore, by Theorem 12.22, $\sin \pi z$ can be expressed as

$$\sin \pi z = e^{g(z)} z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right), \tag{12.10}$$

where $g(z)$ is some entire function. Suppose, for the moment, we could show that $g(z)$ is constant. Then (12.10) could be written as

$$\sin \pi z = cz \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right).$$

Since $\lim_{z\to 0}(\sin \pi z)/z = \pi = \lim_{z\to 0} c \prod_{n=1}^{\infty}(1 - z^2/n^2) = c$, we would then have

$$\sin \pi z = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right). \tag{12.11}$$

The remainder of the proof will consist of showing that $g(z)$ is indeed constant, thus justifying (12.11). To this end, suppose that z is not an integer. Then we form the logarithm derivative in (12.10) to obtain

$$\begin{aligned} \pi \cot \pi z &= g'(z) + \frac{1}{z} + \sum_{n=1}^{\infty} \frac{-2z}{n^2(1 - z^2/n^2)} \qquad (12.12)\\ &= g'(z) + \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2}, \end{aligned}$$

where term-by-term differentiation is justified because $\sum_{n=1}^{\infty} \operatorname{Log}(1 - z^2/n^2)$ converges uniformly on every compact subset of the open set that exclude the integers. It suffices to show that $g'(z) \equiv 0$. To do this, we use the method developed in the proof of Theorem 12.25. Now we consider

$$f(z) = \begin{cases} \pi \cot \pi z - \dfrac{1}{z} & \text{for } z \neq 0 \\ 0 & \text{for } z = 0 \end{cases}$$

and C_n to be the square with vertices at $z = (n + 1/2)(\pm 1 \pm i)$, $n \in \mathbb{N}$. Then $f(z)$ is analytic at the origin having simple poles only at n, $n \in \mathbb{Z}\backslash\{0\}$, and the contour C_n does not pass through the poles of $f(z)$. Clearly, the function $1/z$ is bounded on these squares and for each $n \in \mathbb{Z}\backslash\{0\}$,

$$\operatorname{Res}[f(z); n] = \lim_{z\to n} (z - n) \left[\frac{\pi \cos \pi z}{\sin \pi z} - \frac{1}{z}\right] = \frac{\cos \pi n}{\cos \pi n} = 1.$$

Next we show that there exists a positive number M such that

$$|\cot \pi z| \le M = \frac{1 + e^{-3\pi}}{1 - e^{-3\pi}} \quad \text{for } z \in C_n.$$

To see this, we observe that for $z = n + \frac{1}{2} + iy$ $(y \in \mathbb{R})$,

$$|\cot \pi z| = |\tan(i\pi y)| = \left| \frac{e^{-\pi y} - e^{\pi y}}{e^{-\pi y} + e^{\pi y}} \right| \leq 1 < M.$$

Similarly, for $z = x + i(n + \frac{1}{2})$ $(x \in \mathbb{R})$, we can calculate

$$|\cot \pi z| = \left| \frac{e^{2i\pi z} + 1}{e^{2i\pi z} - 1} \right| \leq \frac{1 + e^{-(2n+1)\pi}}{1 - e^{-(2n+1)\pi}} \leq \frac{1 + e^{-3\pi}}{1 - e^{-3\pi}} = M.$$

Since $\cot(\pi z) = \cot(-\pi z)$, the same bound is valid on the other two sides of C_n, which confirms that $\cot(\pi z)$ is bounded on all contours C_n taken as a whole. By Theorem 12.25, we get

$$\begin{aligned} f(z) &= \lim_{k\to\infty} \sum_{\substack{n=-k \\ n\neq 0}}^{k} \left(\frac{1}{z-n} + \frac{1}{n} \right) \\ &= \lim_{k\to\infty} \sum_{n=1}^{k} \left(\frac{1}{z-n} + \frac{1}{z+n} \right) \\ &= \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2}. \end{aligned}$$

We conclude that the identity

$$\pi \cot \pi z = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2} \tag{12.13}$$

is valid for all nonintegral values of z. A comparison of (12.13) with (12.12) shows that $g'(z) \equiv 0$. But this means that $g(z)$ is constant, which verifies (12.11).

Whenever an infinite product expansion for an entire function is found, it is of interest to compare it with a power series expansion. This can often lead to interesting relationships. To illustrate, we have

$$\sin \pi z = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2} \right) = \pi z - \frac{(\pi z)^3}{3!} + \frac{(\pi z)^5}{5!} - \cdots . \tag{12.14}$$

The z^3 term in the infinite product is

$$\pi z(-z^2) \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots \right) = -\pi z^3 \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

By the uniqueness of the Maclaurin expansion, we must have

$$-\pi \sum_{n=1}^{\infty} \frac{1}{n^2} = -\frac{\pi^3}{3!}.$$

This gives the identity $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$.

As we have seen, the function

$$f(z) = \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n} \tag{12.15}$$

is an entire function having simple zeros at the negative integers and no other zeros. The function $f(z-1)$ has the same zeros in addition to a zero at the origin. Hence

$$f(z-1) = e^{g(z)} z f(z), \tag{12.16}$$

where $g(z)$ is some entire function. Next we show that $g(z)$ is actually a constant. Forming the logarithmic derivative in (12.16), we get

$$\frac{f'(z-1)}{f(z-1)} = g'(z) + \frac{1}{z} + \frac{f'(z)}{f(z)}$$

which, by Theorem 12.24, gives

$$\sum_{n=1}^{\infty} \left(\frac{1}{n+z-1} - \frac{1}{n} \right) = g'(z) + \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{n+z} - \frac{1}{n} \right). \tag{12.17}$$

The sum on the left-hand side of (12.17) can be expressed as

$$\begin{aligned}
&\left(\frac{1}{z} - 1\right) + \sum_{n=2}^{\infty} \left(\frac{1}{n+z-1} - \frac{1}{n} \right) \\
&\qquad = \frac{1}{z} - 1 + \sum_{n=1}^{\infty} \left[\left(\frac{1}{n+z} - \frac{1}{n} \right) + \left(\frac{1}{n} - \frac{1}{n+1} \right) \right] \\
&\qquad = \frac{1}{z} - 1 + \sum_{n=1}^{\infty} \left(\frac{1}{n+z} - \frac{1}{n} \right) + 1 \\
&\qquad = \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{n+z} - \frac{1}{n} \right).
\end{aligned}$$

Comparing this with the right side of (12.17), we find that $g'(z) \equiv 0$. Thus $g(z) = \gamma$, γ a constant. To determine γ, we set $z = 1$ in (12.16). This gives

$$1 = f(0) = e^{\gamma} f(1) = e^{\gamma} \prod_{n=1}^{\infty} \left(1 + \frac{1}{n}\right) e^{-1/n}, \quad \text{or}$$

$$e^{-\gamma} = \lim_{n\to\infty} \left[(1+1)(1+1/2) \cdots (1+1/n) \exp\left(-(1+1/2+ \cdots +1/n)\right) \right].$$

Therefore, using the natural logarithm, we find

$$\begin{aligned}\gamma &= \sum_{n=1}^{\infty}\left[\frac{1}{n} - \ln\left(1+\frac{1}{n}\right)\right] \qquad (12.18)\\ &= \lim_{n\to\infty}\left[1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}-\ln(n+1)\right].\end{aligned}$$

The constant γ in (12.18) is known as *Euler's constant*; its numerical value is approximately 0.577. The fact that this limit exists gives a "sophisticated" way of showing that $\sum_{n=1}^{\infty}(1/n)$ diverges. The function

$$\Gamma(z) = \frac{1}{e^{\gamma z} z f(z)} := \frac{e^{-\gamma z}}{z}\prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)^{-1} e^{z/n} \qquad (12.19)$$

is known as the *gamma function.* It represents an analytic function at all points except the negative integers and zero, where it has simple poles. Since $g(z) = \gamma$ in (12.16), we have

$$f(z) = e^{\gamma}(z+1)f(z+1).$$

This enables us to determine the most important property of the gamma function; namely, the functional equation

$$\Gamma(z+1) = \frac{1}{e^{\gamma(z+1)}(z+1)f(z+1)} = \frac{1}{e^{\gamma z}f(z)} = z\Gamma(z). \qquad (12.20)$$

When $z = n$, a positive integer, (12.20) shows that

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \cdots = n(n-1)\cdots 2\Gamma(1).$$

But

$$\Gamma(1) = \frac{1}{e^{\gamma}f(1)} = \frac{1}{e^{\gamma}e^{-\gamma}} = 1,$$

so that $\Gamma(n+1) = n!$. We collect the above piece of information as

Theorem 12.26. *The gamma function is analytic in $\mathbb{C}$ except at the simple poles at $0, -1, -2, \dots$. Also, $\Gamma(z+1) = \Gamma(z)$ and $\Gamma(n+1) = n!$ for $n \in \mathbb{N}$. Moreover,*

$$\frac{1}{\Gamma(z)} = e^{\gamma z} z \prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)e^{-z/n}$$

is an entire function.

An interesting relationship exists between the gamma function and the sine function. Applying the relations

$$\sin \pi z = \pi z f(z) f(-z) \quad \text{and} \quad f(z) = 1/z e^{\gamma z}\Gamma(z)$$

in (12.19), we obtain

$$\sin \pi z = \frac{\pi}{e^{\gamma z}\Gamma(z)(-z)e^{-\gamma z}\Gamma(-z)} = \frac{\pi}{-z\Gamma(-z)\Gamma(z)}.$$

Since $-z\Gamma(-z) = \Gamma(1-z)$, it follows that

$$\frac{\pi}{\sin \pi z} = \Gamma(z)\Gamma(1-z)$$

for all nonintegral values of z. In particular, setting $z = \frac{1}{2}$, we find that $\Gamma^2\left(\frac{1}{2}\right) = \pi$. Because Γ is obviously a positive function on $(0, \infty)$, we deduce that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. Also, it follows that $\Gamma(z)$ is zero-free because for $z \neq 0, -1, -2, \dots$, the gamma function is given by a convergent infinite product of nonvanishing factors. Further, for $n \in \mathbb{N}$, the functional equation gives

$$\Gamma\left(\frac{2n+1}{2}\right) = \left(\frac{2n-1}{2}\right)\Gamma\left(\frac{2n-1}{2}\right) = \frac{1\cdot 3\cdots(2n-1)}{2^n}\Gamma\left(\frac{1}{2}\right),$$

whereas the above identity gives

$$\Gamma\left(-\frac{2n+1}{2}\right) = \frac{\pi}{\sin \pi(n+\frac{3}{2})}\frac{1}{\Gamma\left(\frac{2n+3}{2}\right)} = \frac{(-1)^{n+1}\pi}{\Gamma\left(\frac{2n+3}{2}\right)} = \frac{(-1)^{n+1}2^{n+1}\sqrt{\pi}}{1\cdot 3\cdots(2n+1)}.$$

Remark 12.27. In real analysis, the function

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt \quad (x > 0)$$

is studied extensively. Interestingly enough, the gamma function defined by (12.19) can be expressed in this integral form for all positive real values of z. In the next chapter, we will redefine the gamma function as a complex integral and show that the two definitions represent the same function at all points where the integral converges. ●

Remark 12.28. We may also evaluate $\Gamma(\frac{1}{2})$ by this integral definition. We have

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-1/2}e^{-t}\,dt.$$

A substitution of $t = y^2$ leads to

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^\infty e^{-y^2}\,dy = 2\frac{\sqrt{\pi}}{2} = \sqrt{\pi},$$

this last integral having been evaluated in Section 9.3. ●

Questions 12.29.

1. Can an entire function be constructed that has "a" points at a preassigned sequence of points?
2. If $a_n \to a \in \mathbb{C}$ and if f is entire with zeros at a_n, then is $f(z) \equiv 0$ in $\mathbb{C}$?

3. If $\sum_{n=1}^{\infty} 1/a_n$ diverges, can $\prod_{n=1}^{\infty}(1 - z/a_n)$ be analytic anywhere?
4. What is meant by a "convergence producing" factor?
5. Is there an analog to Weierstrass's theorem for functions analytic in an arbitrary domain?
6. What other choices for the sequence $\{P_n(z)\}$ would have worked in the proof of Weierstrass's theorem?
7. If $\sum_{n=1}^{\infty} a_n$ converges, is $\prod_{n=1}^{\infty}(1+a_n z)$ entire? Is $\prod_{n=1}^{\infty}(1+a_n z^2)$ entire? Is $\prod_{n=1}^{\infty}(1 + a_n p(z))$ entire? Here $p(z)$ is a polynomial.
8. What is the relationship between the infinite product and infinite series expansion for entire functions?
9. Knowing the infinite product expansion for $\sin \pi z$, what other infinite product expansions can we determine?

Exercises 12.30.

1. Construct an analytic function f in $|z| < R$ such that f has zeros only at $z = -R + 1/n$, $n \in \mathbb{N}$.
2. Construct an entire function whose only zeros are at $z = \ln n$ ($n = 2, 3, 4, \dots$).
3. Construct an entire function $f(z)$ with the following properties:
   a) $f(z)$ vanishes at $z = 1, 2, 3, \dots$ and nowhere else.
   b) The zero of $f(z)$ at $z = n$ has multiplicity n.
   c) Construct an entire function $f(z)$ such that f has zeros at $z = n^{3/2}$ ($n \in \mathbb{N}$) and nowhere else.
   d) Construct an entire function $f(z)$ such that f has zeros at $z = n^{3/4}$ ($n \in \mathbb{N}$) and nowhere else.
4. (a) Find the value of $\prod_{n=1}^{\infty}(1 + 1/n^2)$.
   (b) Show that $\prod_{n=2}^{\infty}(1 - 1/n^4) = (e^{\pi} - e^{-\pi})/8\pi$.
5. (a) Show that $e^{2z} - 1$ and $\sin iz$ have simple zeros at the same points.
   (b) Set $(e^{2z} - 1)/\sin iz = e^{g(z)}$, and determine $g(z)$.
6. By comparing the term involving z^5 for the series and the product expansion of $\sin \pi z$, show that $\sum_{n=1}^{\infty} n^{-4} = \pi^4/90$.
7. Use the summation formula for $\pi \cot \pi z$ to sum the series $\sum_{n=1}^{\infty} n^{-2}$ and $\sum_{n=1}^{\infty} n^{-4}$.
8. Show that the convergence of (12.13) is uniform on all compact subsets that contain no integers.
9. Show that
$$\prod_{n=1}^{\infty} \left(1 - \frac{z}{n}\right) e^{z/n+(5z^2/n^2)}$$
represents an entire function.
10. Suppose $0 < |a_1| \le |a_2| \le \cdots \to \infty$. Show that $\prod_{n=1}^{\infty} \left(1 - \frac{z}{a_n}\right) e^{Q_n(z)}$ represents an entire function, where
$$Q_n(z) = \frac{z}{a_n} + \frac{1}{2}\left(\frac{z}{a_n}\right)^2 + \cdots + \frac{1}{[\ln n]}\left(\frac{z}{a_n}\right)^{[\ln n]}.$$

11. Prove that

$$\cos z = \prod_{n=1}^{\infty} \left[1 - \frac{4z^2}{(2n-1)^2\pi^2}\right].$$

12. Prove that

$$\pi \coth \pi z = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 + n^2}.$$

13. If a is not an integer, show that

$$f(z) = \prod_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \left(1 + \frac{z}{a+n}\right) e^{-z/(n+a)}$$

is an entire function, and that

$$f(z) = \frac{a \sin \pi(z+a)}{(z+a)\sin \pi a} e^{-\pi z \cot(\pi a) e^{z/a}}.$$

14. Evaluate $\prod_{n=1}^{\infty}(1 + 1/n^2 + 1/n^4)$.
15. Use the product expansion for $\sin \pi z$ to show that
    (a) $\dfrac{\pi}{2} = \left(\dfrac{2\cdot 2}{1\cdot 3}\right)\left(\dfrac{4\cdot 4}{3\cdot 5}\right)\left(\dfrac{6\cdot 6}{5\cdot 7}\right)\cdots.$
    (b) $\sqrt{2} = \left(\dfrac{2\cdot 2}{1\cdot 3}\right)\left(\dfrac{6\cdot 6}{5\cdot 7}\right)\left(\dfrac{10\cdot 10}{9\cdot 11}\right)\cdots.$
    (c) $\sqrt{3} = 2\left(\dfrac{2\cdot 4}{3\cdot 3}\right)\left(\dfrac{8\cdot 10}{9\cdot 9}\right)\left(\dfrac{14\cdot 16}{15\cdot 15}\right)\cdots.$
16. Show that

$$\frac{d}{dz}\left(\frac{\Gamma'(z)}{\Gamma(z)}\right) = \sum_{n=0}^{\infty} \frac{1}{(z+n)^2},$$

where $\Gamma(z)$ is defined by (12.19). Here the function $\Psi = \Gamma'/\Gamma$ is known as the digamma function, or alternatively, as the Gauss psi-function.
17. Show that $\Gamma(z)\sin \pi z$ is an entire function.
18. Expand $e^z - 1$ in an infinite product.
19. Construct an entire function having simple zeros at $z = n^2$ $(n \in \mathbb{N})$ and nowhere else. Show that one solution to this is given by the entire function $(\sin \pi\sqrt{z})/(\pi\sqrt{z})$.

12.3 Mittag-Leffler Theorem

A function is said to be *meromorphic* in a domain D if it has no singularities, other than possibly poles, in D. If no domain is specified, it will be assumed that the function is meromorphic in the whole complex plane. Thus, every entire function is meromorphic but the converse is not necessarily true. Sum and products of meromorphic functions are meromorphic. Quotients of meromorphic functions are meromorphic, provided that the denominator is not identically zero. For $b \in \mathbb{C}$, the function $1/(z-b)$ is an example of the simplest type of meromorphic function. As a consequence of the definition, we see that the function $f(z)$ which is meromorphic in $\mathbb{C}$ cannot have infinitely many poles in a bounded region. For if it does, the sequence of poles must have a limit point p. Since $f(z)$ cannot be analytic in any neighborhood of p, the point p is a singularity that is not a pole. Note that

$$\frac{1}{\sin(1/(1-z))}$$

has infinitely many poles in the unit disk $|z| < 1$.

Example 12.31. From Theorem 12.26, we observe that the gamma function defined by (12.19) is a meromorphic function in $\mathbb{C}$ with simple poles at $0, -1, -2, \dots$ and the functional equation $\Gamma(z+1) = \Gamma(z)$ shows that

$$\Gamma(z+n) = (z+n-1)\Gamma(z+n-1) = (z+n-1)\,\cdots\,(z+1)z\Gamma(z),$$

for $z \neq 0, -1, -2, \dots$. Using this, we see that

$$\begin{aligned}\lim_{z\to -n}(z+n)\Gamma(z) &= \lim_{z\to -n}\frac{\Gamma(z+n+1)}{(z+n-1)\,\cdots\,(z+1)z}\\ &= \frac{\Gamma(1)}{(-1)(-2)\,\cdots\,(-n+1)(-n)} = \frac{(-1)^n}{n!},\end{aligned}$$

and therefore $\mathrm{Res}\,[\Gamma(z); -n] = (-1)^n/n!$. ●

The reciprocal of an entire function is meromorphic, its poles consisting of the zeros of the entire function. More generally, we have the following characterization of meromorphic functions.

Theorem 12.32. *A function is meromorphic if and only if it can be expressed as the quotient of entire functions.*

Proof. First, suppose that $f(z) = g(z)/h(z)$, where $g(z)$ and $h(z)$ are entire functions with no common zeros (any common zero can be factored out). Then the only singularities of $f(z)$ are poles consisting of the zeros of $h(z)$. Hence $f(z)$ is meromorphic.

Conversely, suppose $f(z)$ is meromorphic. Then by Weierstrass's theorem, there exists a function $h(z)$ where zeros coincide in both position and order with the poles of $f(z)$. Therefore, $g(z) = f(z)h(z)$ is an entire function because the poles of $f(z)$ are cancelled by the zeros of $h(z)$. Thus $f(z) = g(z)/h(z)$ is the quotient of entire functions, and the proof is complete. ■

In this section, we are going to prove a theorem for meromorphic functions analogous to Weierstrass's theorem for entire functions. Given any finite set of points $\{b_1, b_2, \dots, b_n\}$, the (meromorphic) rational function

$$\frac{1}{z-b_1}+\frac{1}{z-b_2}+\cdots+\frac{1}{z-b_n}$$

has a simple pole at each point of the set. Therefore, any other meromorphic function $f(z)$ having a simple pole at b_k with the principal part as above must be of the form

$$f(z)=\sum_{k=1}^{n}\frac{1}{z-b_k}+\phi(z)$$

where $\phi(z)$ is an arbitrary entire function. If the given set is infinite, the problem of constructing a meromorphic function having a pole at each point of the set is more complicated. Then we have to worry about the convergence.

Consider the function

$$f(z)=\sum_{n=1}^{\infty}\frac{1}{z^2-n^2},$$

which is defined and converges for all values of z except for the squares of integers. For all z in some neighborhood of a given point not containing the square of an integer, we have

$$\frac{1}{|z^2-n^2|}\leq\frac{2}{n^2}$$

for n sufficiently large. Hence by Theorem 6.31, the convergence of the series for $f(z)$ is uniform in some neighborhood of each such point. This shows that $f(z)$ is analytic at all points except $z=n^2$ $(n\in\mathbb{N})$ so that $f(z)$ is a meromorphic function having simple pole at $z=n^2$.

On the other hand, suppose we wish to construct a meromorphic function having simple poles at the positive integers with residues equal to 1. The likely candidate $f(z)=\sum_{n=1}^{\infty}1/(z-n)$ fails to converge anywhere. As in the case of Weierstrass theorem, a convergence producing term is needed. Note that the constant term in the power series expansion of $1/(z-n)$ about $z=0$ is $-1/n$. So we try with the function

$$\sum_{n=1}^{\infty}\left(\frac{1}{z-n}+\frac{1}{n}\right)=\sum_{n=1}^{\infty}\frac{z}{n(z-n)}.$$

Note that for $|z|\leq R$ and N large enough so that $N\geq 2R$, then

$$|z-n|\geq n-|z|\geq n-R\geq n/2 \quad\text{for all } n\geq N.$$

Therefore,

$$\left|\frac{1}{n(z-n)}\right|\leq\frac{2}{n^2}\quad\text{for}\quad n\geq N$$

so that the series converges absolutely for all z excluding positive integers.

In one sense, the theorem we prove about meromorphic functions will be more general than the corresponding theorem for entire functions. For we will construct a meromorphic function not only with preassigned poles, but with preassigned principal parts. Let $\{b_n\}$ be a sequence tending to ∞. For each n, associate a rational function of the form

$$P_n\left(\frac{1}{z-b_n}\right) = \frac{a_1^{(n)}}{z-b_n} + \frac{a_2^{(n)}}{(z-b_n)^2} + \cdots + \frac{a_{k_n}^{(n)}}{(z-b_n)^{k_n}}, \qquad (12.21)$$

where the $a_i^{(n)}$ are complex constants, $a_{k_n}^{(n)} \neq 0$. Note that k_n is the order of the pole, and may vary with n. Our goal is to construct a meromorphic function whose principal part for each n is $P_n(1/(z-b_n))$. If the series $\sum_{n=1}^{\infty} P_n(1/(z-b_n))$ does not converge, it will be shown that the convergence producing factor for each n consists of a polynomial that is the partial sum of the Maclaurin expansion for $P_n(1/(z-b_n))$.

Before stating and proving our theorem, one final remark is in order. Suppose two meromorphic functions $f(z)$ and $g(z)$ have the same poles with the same principal parts. Then $f(z) - g(z)$ has no poles, and consequently must be an entire function. Hence any two meromorphic functions with the same principal parts can differ by at most an (additive) entire function. This entire function is the meromorphic analog to the (multiplicative) exponential function of the corollary to Theorem 12.22. Now we formulate this discussion as

Theorem 12.33. *Let $f(z)$ be a meromorphic function. If $\phi(z)$ is an arbitrary entire function, then the most general meromorphic function $g(z)$ which coincides with $f(z)$ in its poles and the corresponding principal parts is given by $g(z) = f(z) + \phi(z)$.*

We now illustrate this result by an example. Consider

$$f(z) = \cot z \quad \text{and} \quad g(z) = \frac{2ie^{2iz}}{e^{2iz}-1}.$$

Then both $f(z)$ and $g(z)$ are meromorphic functions in $\mathbb{C}$ having simple poles at $z = n\pi$, $n \in \mathbb{Z}$. It is a simple exercise to see that, for both the functions, residues at each of these poles are 1. Thus, the poles and the corresponding principal parts of f and g are the same. Consequently, they differ by an additive entire function. Again, it is easy to see that $f(z) - g(z) = -i$, an entire function.

The dominating result for meromorphic functions in $\mathbb{C}$ is due to Mittag-Leffler.

Theorem 12.34. (Mittag-Leffler's Theorem) *Let $\{b_n\}_{n\geq 1}$ be a sequence of points tending to ∞, and $P_n(1/(z-b_n))$ be a polynomial in $1/(z-b_n)$ of the form* (12.21). *Then there exists a meromorphic function $f(z)$ that has poles at the points b_n ($n \in \mathbb{N}$) with principal part $P_n(1/(z-b_n))$, and is otherwise analytic.*

Proof. Without loss of generality, assume that none of the b_n is zero, for we can always add a rational function having a pole at the origin. It may also be assumed that the sequence is so arranged that

$$0 < |b_1| \le |b_2| \le |b_3| \le \cdots .$$

For each n, the rational function $P_n(1/(z-b_n))$ is analytic in the disk $|z| < |b_n|$ and has a Maclaurin expansion

$$P_n\left(\frac{1}{z-b_n}\right) = \sum_{k=0}^{\infty} a_k^{(n)} z^k.$$

This series clearly converges absolutely in $|z| < |b_n|$ and uniformly in $|z| \le |b_n|/2$. Thus, $P_n(1/(z-b_n))$ can be approximated in $|z| \le |b_n|/2$ by a partial sum

$$Q_n(z) = \sum_{k=0}^{n_k} a_k^{(n)} z^k = a_0^{(n)} + a_1^{(n)} z + \cdots + a_{n_k}^{(n)} z^{n_k}$$

as closely as we please. In particular, for a large value of n, we have

$$\left| P_n\left(\frac{1}{z-b_n}\right) - Q_n(z) \right| < \frac{1}{2^n} \quad \text{for} \quad |z| \le \frac{|b_n|}{2}. \tag{12.22}$$

We will show that the function

$$f(z) = \sum_{n=1}^{\infty} \left[P_n\left(\frac{1}{z-b_n}\right) - Q_n(z) \right]$$

is the meromorphic function that we want. It suffices to show that the series $f(z)$ converges uniformly on an arbitrary compact subset $|z| \le R$ of $\mathbb{C}$ that excludes the points $|b_n| \le R$. Choosing N so that $|b_N| \ge 2R$, we see from (12.22) that

$$\sum_{n=N}^{\infty} \left| P_n\left(\frac{1}{z-b_n}\right) - Q_n(z) \right| < \sum_{n=N}^{\infty} \frac{1}{2^n} < \infty.$$

The uniform convergence follows from the Weierstrass M-test (see Theorem 6.31). Note that

$$\sum_{n=N}^{\infty} \left[P_n\left(\frac{1}{z-b_n}\right) - Q_n(z) \right]$$

is analytic in $|z| \le R$ because the poles of $P_n(1/(z-b_n))$ lie outside $|z| = R$. For $|z| \le R$ $(\le |b_N|/2)$

$$\sum_{n=0}^{N-1} \left[P_n\left(\frac{1}{z-b_n}\right) - Q_n(z) \right]$$

is an analytic function with no singularities except the prescribed poles. Since R is arbitrary, $f(z)$ is meromorphic in the plane. ∎

Remark 12.35. Instead of choosing the convergence producing polynomials $\{Q_n(z)\}$, we could have chosen any other sequence of polynomials $\{R_n(z)\}$ for which $\sum_{n=1}^{\infty}[P_n(z) - R_n(z)]$ converges uniformly on compact subsets that exclude the poles. ●

Remark 12.36. In general, the degree of the convergence producing polynomials $\{R_n(z)\}$ varies with n. If $\sum_{n=1}^{\infty} P_n(1/(z-b_n))$ converges, we may choose $R_n(z) \equiv 0$. If

$$P_n\left(\frac{1}{z-b_n}\right) = \frac{1}{z-n} = -\frac{1}{n}\sum_{k=0}^{\infty}\left(\frac{z}{n}\right)^k \quad (|z| < n),$$

it was shown that we may choose $R_n(z) = 1/n$, a sequence of constant polynomials. The reader may also verify that if

$$P_n\left(\frac{1}{z-b_n}\right) = \frac{1}{z-\sqrt{n}} = -\frac{1}{\sqrt{n}}\sum_{k=0}^{\infty}\frac{z^k}{n^{k/2}} \quad (|z| < \sqrt{n}),$$

we may choose $R_n(z) = 1/\sqrt{n} + z/n$. ●

The function

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

is seen to be a meromorphic function having a simple pole at each integer with residue 1. Hence

$$f(z) = \frac{1}{z} + \sum_{n=1}^{\infty}\frac{2z}{z^2-n^2} + g(z), \tag{12.23}$$

where $g(z)$ is entire, is the most general such meromorphic function. In the special case that $g(z) \equiv 0$, a comparison of (12.23) and (12.13) shows that $f(z) = \pi\cot\pi z$.

Throughout this section, we have seen similarities between Weierstrass's theorem and Mittag-Leffler's theorem. As an application of Mittag-Leffler's theorem, we will now prove a generalization of Weierstrass's theorem.

Theorem 12.37. *Let $\{a_n\}$ be the sequence of distinct complex numbers approaching ∞. Then for any sequence $\{c_n\}$ of complex numbers, there exists an entire function $f(z)$ such that $f(a_n) = c_n$ for every n.*

Proof. According to Weierstrass's theorem we construct an entire function $g(z)$ that has simple zeros at $z = a_n$. Then $g(a_n) = 0$ and $g'(a_n) \neq 0$ for each n. According to Mittag-Leffler's theorem, we can construct a meromorphic function $h(z)$ that has simple poles at $z = a_n$ with the principal part $c_n/g'(a_n)(z-a_n)$ (if $c_n = 0$, $h(z)$ is taken to be analytic at $z = a_n$).

Since the simple poles of $h(z)$ are also the simple zeros of $g(z)$, the singularities of $f(z) = g(z)h(z)$ are removable. That is, $f(z)$ is an entire function. For each n, we can expand $g(z)$ in a Taylor series about the point $z = a_n$. Then

$$g(z) = g'(a_n)(z - a_n) + \frac{g''(a_n)}{2}(z - a_n)^2 \cdots . \tag{12.24}$$

Also we may write

$$h(z) = \frac{c_n}{g'(a_n)(z - a_n)} + h_1(z), \tag{12.25}$$

where $h_1(z)$ is analytic in some neighborhood of $z = a_n$. Combining (12.24) and (12.25), we find

$$f(a_n) = \lim_{z \to a_n} f(z) = \lim_{z \to a_n} g(z)h(z) = c_n.$$

This completes the proof. ∎

Remark 12.38. If $c_n \equiv 0$, then Theorem 12.37 reduces to Weierstrass's theorem. ●

Questions 12.39.

1. What is the relationship between Weierstrass's theorem and Mittag-Leffler's theorem? Can one be derived from the other?
2. What can be said about the sum of meromorphic functions? The product?
3. How do the convergence producing factors of Weierstrass's theorem and Mittag-Leffler's theorem compare?
4. Why is the logarithmic derivative important in this chapter?
5. In the proof of Mittag-Leffler's theorem, why was it necessary to assume that none of the b_n were equal to zero?
6. Is there a unique entire function that satisfies the conditions of Theorem 12.37?
7. Is Mittag-Leffler's theorem still valid if we allow the principal part to have essential singularities?

Exercises 12.40.

1. Construct a meromorphic function $f(z)$ with the following two properties:
   (i) $f(z)$ has poles at $z = 1, 2, 3, \ldots$ and nowhere else.
   (ii) The pole at $z = n$ has order n.
2. Show that $(\sin z)/(e^{2iz} + 1)$ and $1/(2i \cos z)$ differ by an entire function, and determine it.
3. Show that $\prod_{n=1}^{\infty}(1 - z/a_n)$ is entire if and only if $\sum_{n=1}^{\infty} 1/(z - a_n)$ is meromorphic.

4. Let $(p \geq 1)$ be an integer. If $\sum_{n=1}^{\infty} 1/|a_n|^{p+1}$ converges, show that

$$\sum_{n=1}^{\infty} \left(\frac{1}{z-a_n} + \frac{1}{a_n} + \frac{z}{a_n^2} + \frac{z^2}{a_n^3} + \cdots + \frac{z^{p-1}}{a_n^p} \right)$$

is meromorphic.

5. (a) Suppose that $0 < |b_1| \leq |b_2| \leq \cdots$ $(|b_n| \to \infty)$. Show that there exists a positive sequence $\{a_n\}$ such that $f(z) = \sum_{n=1}^{\infty} a_n/(z-b_n)$ is analytic except at $z = b_n$.
   (b) Consider the expansion $f(z) = c_0 + c_1 z + c_2 z^2 + \cdots + c_m z^m + \cdots$ and express c_m in terms of the quantities a_n and b_n.
6. Suppose $g(z) = \prod_{n=1}^{\infty}(1 - z/b_n)$ is entire. With the notation of the previous exercise, show that $h(z) = f(z)g(z)$ is entire and evaluate $h(z_n)$ in terms and α_n and $g'(z_n)$.
7. Suppose $\sum_{n=1}^{\infty}(|\alpha_n|/n^2)$ converges. Show that
   (a) $f(z) = 2z\sum_{n=1}^{\infty}(-1)^n \alpha_n/(z^2 - n^2)$ is meromorphic.
   (b) $g(z) = (\sin \pi z/\pi) f(z)$ is entire, with $g(n) = \alpha_n$.
8. Show that

$$\frac{\pi^2}{\sin^2(\pi z)} = \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{(z-n)^2}.$$

9. Derive the Weierstrass's theorem from Mittag-Leffler's theorem.
10. Prove the identity

$$\frac{1}{e^z - 1} = \frac{1}{z} - \frac{1}{2} + 2z \sum_{n=1}^{\infty} \frac{1}{z^2 + 4\pi^2 n^2}.$$

11. Show that the function $\Theta(z)$ defined by

$$\Theta(z) = \prod_{k=1}^{\infty} \left(1 + h^{2k-1} e^z\right)\left(1 + h^{2k-1} e^{-z}\right) \quad (0 < |h| < 1),$$

is entire and satisfies the functional equation

$$\Theta(z + 2\log h) = h^{-1} e^{-z} \Theta(z).$$

13

Analytic Continuation

We have previously seen that an analytic function is determined by its behavior at a sequence of points having a limit point. This was precisely the content of the identity theorem (see Theorem 8.48) which is also referred to as the principle of analytic continuation. For example, as a consequence, there is precisely a unique entire function on $\mathbb{C}$ which agrees with $\sin x$ on the real axis, namely $\sin z$. But we have not yet explored the following question: If $f(z)$ is analytic in a domain D_1, is there a function analytic in a different domain D_2 that agrees with $f(z)$ in $D_1 \cap D_2$? Analytic continuation deals with the problem of properly redefining an analytic function so as to extend its domain of analyticity. In the process, we come across functions for which no such extension exists. Finally, we apply our knowledge of analytic continuation to two of the most important functions in analysis, the gamma function and the Riemann-zeta function, defined originally by a definite integral and an infinite series, respectively.

13.1 Basic Concepts

Consider the power series

$$f_0(z) = \sum_{n=0}^{\infty} z^n.$$

This power series converges for $|z| < 1$, and hence, $f_0(z)$ is analytic in the disk $|z| < 1$ and represents there the function $f(z) = 1/(1-z)$. Although the power series diverges at each point on $|z| = 1$, $f(z)$ is analytic in $\mathbb{C} \setminus \{1\}$. For any point $z_0 \neq 1$, the Taylor series representation

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n \tag{13.1}$$

is valid when $|z - z_0| < |1 - z_0|$ (see Figure 13.1). The disk in which (13.1)

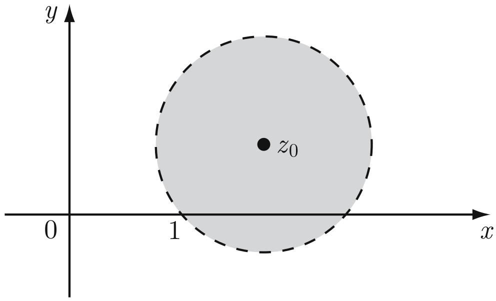


Figure 13.1.

converges may or may not have points in common with the disk $|z| < 1$. For example,

$$f_1(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(e^{i\alpha})}{n!}(z - e^{i\alpha})^n \quad (0 < \alpha < 2\pi)$$

converges in a disk that overlaps $|z| < 1$; but the disk, $|z - 2| < 1$, in which

$$f_2(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(2)}{n!}(z - 2)^n$$

converges does not. In Figure 13.2, we show the domains in which $f_0(z)$, $f_1(z)$, and $f_2(z)$ converge. In their respective domains of convergence, they all represent the same function $f(z) = 1/(1 - z)$. In addition, the integral

$$\int_0^{\infty} e^{-t(1-z)}\, dt$$

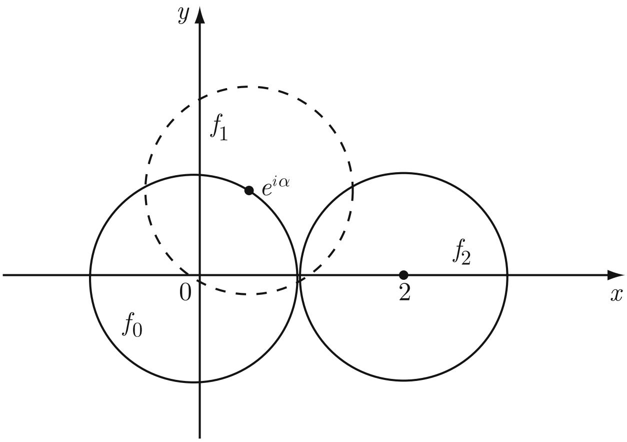


Figure 13.2.

converges for $\operatorname{Re} z < 1$ and it can be easily checked that the integral represents $f(z) = 1/(1-z)$ in this half-plane. But they agree with $f_0(z) = \sum_{n=0}^{\infty} z^n$ ($|z| < 1$) for a certain value of z although they appear different. In fact they agree with $f(z) = 1/(1-z)$ which is analytic for all $z \neq 1$. So we see that apparently unrelated functions may actually represent the same analytic function in different domains.

Suppose $f_0(z)$ is known to be analytic in a domain D_0. We wish to determine the largest domain $D \supset D_0$ for which there exists an analytic function $f(z)$ such that $f(z) \equiv f_0(z)$ in D_0. As we have just seen in the first example, $\mathbb{C}\backslash\{1\}$ is the largest domain containing $|z| < 1$ in which an analytic function may be defined that agrees with $f_0(z) = \sum_{n=0}^{\infty} z^n$ in $|z| < 1$. In our terminology, we say that f_0 has an analytic continuation from the unit disk $|z| < 1$ into the punctured plane $\mathbb{C}\backslash\{1\}$. To see how one can carry out the process of analytic continuations, we need to introduce several definitions.

A function $f(z)$, together with a domain D in which it is analytic, is said to be a *function element* and is denoted by (f, D). Two function elements (f_1, D_1) and (f_2, D_2) are called *direct analytic continuations* of each other iff

$$D_1 \cap D_2 \neq \emptyset \quad \text{and} \quad f_1 = f_2 \quad \text{on } D_1 \cap D_2.$$

Whenever there exists a direct analytic continuation of (f_1, D_1) into a domain D_2, it must be uniquely determined, for any two direct analytic continuations would have to agree on $D_1 \cap D_2$, and by the identity theorem (see Theorem 8.48) would consequently have to agree throughout D_2. That is, given an analytic function f_1 on D_1, there is at most one way to extend f_1 from D_1 into D_2 so that the extended function is analytic in D_2. Thus, one of the main uses of this idea is to extend the functional relations, initially valid for a small domain D_1, to a larger domain D_2. Sometimes such an extension may not be possible. For instance, if D_1 is the punctured unit disk $0 < |z| < 1$ and D_2 is the unit disk, then the function $f_1(z) = 1/z$ cannot be extendable analytically from D_1 into D_2. Similarly, if

$$D_1 = \mathbb{C}\backslash\{z : \operatorname{Re} z \leq 0, \operatorname{Im} z = 0\}, \quad \text{and} \quad D_2 = \mathbb{C},$$

then, for $f_1(z) = \operatorname{Log} z$, no extension from D_1 to D_2 is possible.

Remark 13.1. Consider the series

$$f_1(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}.$$

This series converges for $|z| \leq 1$ and $f_1(z)$ is analytic in the disk $|z| < 1$, and represents the function

$$f(z) = -\int_0^z \frac{\operatorname{Log}(1-t)}{t}\, dt.$$

However, $f_1(z)$ cannot be continued analytically to a domain D with $1 \in D$, since

$$f_1''(z) = \sum_{n=2}^{\infty} \frac{n-1}{n} z^{n-2} \longrightarrow \infty \quad \text{as } z \to 1^+.$$

This observation shows that the convergence or divergence of power series at a point on the circle of convergence does not determine whether the function which defines the series can or cannot be continued along that point. ●

The property of being a direct analytic continuation is not transitive. That is, even if (f_1, D_1) and (f_2, D_2) are direct analytic continuations of each other, and (f_2, D_2) and (f_3, D_3) are direct analytic continuations of each other, we cannot conclude that (f_1, D_1) and (f_3, D_3) are direct analytic continuations of each other. A simple example of this occurs whenever D_1 and D_3 have no points in common. However, there is a relationship between $f_1(z)$ and $f_3(z)$ that is worth exploring.

Suppose $\{(f_1, D_1), (f_2, D_2), \ldots, (f_n, D_n)\}$ is a finite set of function elements with the property that (f_k, D_k) and (f_{k+1}, D_{k+1}) are direct analytic continuations of each other for $k = 1, 2, 3, \ldots, n-1$. Then the set of function elements are said to be *analytic continuations* of one another. Such a set of function elements is then called a *chain*.

Example 13.2. Define (see Figure 13.3)

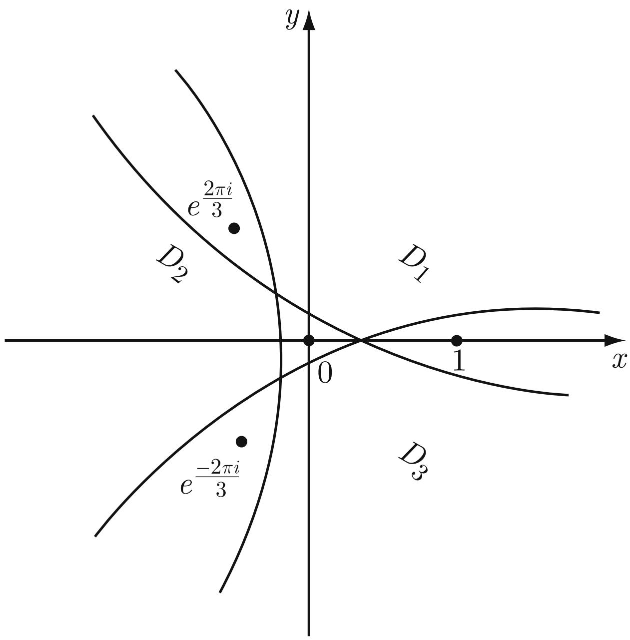


Figure 13.3. Illustration for a chain with $n = 3$

$$\begin{aligned} f_1(z) &= \operatorname{Log} z \quad \text{for } z \in D_1 \\ f_2(z) &= \operatorname{Log} z \quad \text{for } z \in D_2 \\ f_3(z) &= \operatorname{Log} z + 2\pi i \quad \text{for } z \in D_3. \end{aligned}$$

Then $\{(f_1, D_1), (f_2, D_2), (f_3, D_3)\}$ is a chain with $n = 3$. Note that $0 = f_1(1) \neq f_3(1) = 2\pi i$. ●

Note that (f_i, D_i) and (f_j, D_j) are analytic continuations of each other if and only if they can be connected by finitely many direct analytic continuations. If $\gamma : [0,1] \to \mathbb{C}$ is a curve and if there exists a chain $\{(f_i, D_i)\}_{1\leq i\leq n}$, of function elements such that

$$\gamma([0,1]) \subset \cup_{i=1}^{n} D_i,\ z_0 = \gamma(0) \in D_1,\ z_n = \gamma(1) \in D_n,$$

then we say that the function element (f_n, D_n) is an analytic continuation of (f_1, D_1) along the curve γ. That is a function element (f, D) can be analytically continued along a curve if there is a chain containing (f, D) such that each point on the curve is contained in the domain of some function element of the chain. As another example, the domains of a chain are also shown in Figure 13.4. In some situations, analytic continuation of function element are carried out easily by means of power series. In this case, a chain is a sequence of overlapping disks.

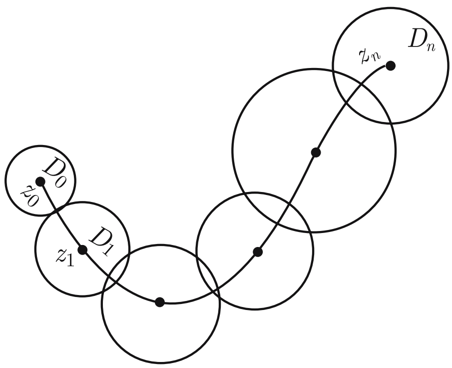


Figure 13.4. Illustration for a chain

Given a chain $\{(f_1, D_1), (f_2, D_2), \dots, (f_n, D_n)\}$, can a function $f(z)$ be defined such that $f(z)$ is analytic in the domain $\{D_1 \cup D_2 \cup \cdots \cup D_n\}$? Certainly this can be done when $n = 2$. The function

$$f(z) = \begin{cases} f_1(z) & \text{if } z \in D_1 \\ f_2(z) & \text{if } z \in D_2, \end{cases}$$

is analytic in $D_1 \cup D_2$. If $D_1 \cap D_2 \cap \cdots \cap D_n \neq \emptyset$, we can show by induction that f defined by $f(z) = f_i(z)$ for $z \in D_i$ $(i = 1, 2, \dots, n)$ is analytic. However,

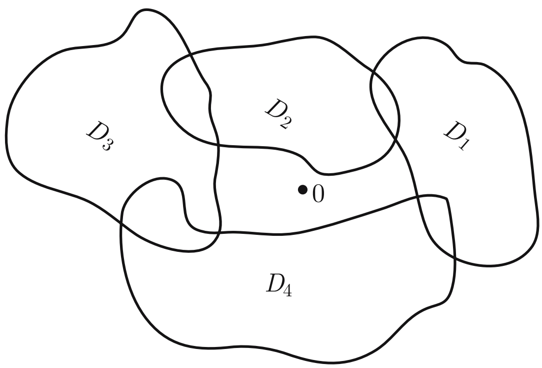


Figure 13.5.

the proof for the general case fails. Consider the four domains illustrated in Figure 13.5. For a fixed branch of $\log z$, set $f_1(z) = \log z$ in D_1. The function element (f_1, D_1) determines a unique direct analytic continuation (f_2, D_2), which determines (f_3, D_3), which determines (f_4, D_4). We thus have the chain $\{(f_1, D_1), (f_2, D_2), (f_3, D_3), (f_4, D_4)\}$. However, in the domain $D_1 \cap D_4$ it is not true that $f_1(z) = f_4(z)$. We actually have $f_4(z) = f_1(z) + 2\pi i$ for all points in $D_1 \cap D_4$. The difference in the two functions lies in the fact that the argument of the multiple-valued logarithmic function has increased by 2π after making a complete revolution around the origin. Note also that we can continue (f_1, D_1) into the domain D_3 by different chains and come up with different functions. For the chains $\{(f_1, D_1), (f_2, D_2), (f_3, D_3)\}$ and $\{(f_1, D_1), (g_1, D_4), (g_2, D_3)\}$, we have the values of f_3 and g_2 differing by $2\pi i$. Before we continue the discussion, let us present our case by a concrete example.

Example 13.3. Consider the function $f(z)$, initially defined on the disk $D = \{z : |z - 1| < 1\}$ by the series expansion

$$f(z) = z^{1/2} = 1 + \frac{1}{2}(z-1) - \frac{1}{8}(z-1)^2 + \cdots .$$

Here it is understood that we start with the series representation of the principal branch of $\sqrt{z}$:

$$f(z) = e^{(1/2)\,\mathrm{Log}\, z} = (1 + (z-1))^{1/2}.$$

Note also that f is analytic in D. Let $\gamma : [0, 2\pi] \to \mathbb{C}$ be the closed contour given by $\gamma(t) = e^{it}$, starting from $z_0 = \gamma(0) = 1$. Then $f(z)$ actually has an analytic continuation along γ. In fact, we have an explicit convergent power series about e^{it} (write $z^{1/2} = e^{it/2}[1 + (z - e^{it})/e^{it}]^{1/2}$):

$$f_t(z) = e^{it/2} + \frac{1}{2}e^{-it/2}(z - e^{it}) + \frac{1}{2}\left(\frac{1}{2} - 1\right)\frac{1}{2!}e^{-3it/2}(z - e^{it})^2 + \cdots ,$$

where $z \in D_t = \{z : |z - e^{it}| < 1\}$. Thus, after one complete round along the unit circle, we end up at $z = 2\pi$ by

$$f_{2\pi}(z) = -\left[1 + \frac{1}{2}(z-1) - \frac{1}{8}(z-1)^2 + \cdots\right]$$

which is just the other branch of $\sqrt{z}$. The initial and final function elements in this case are $(e^{(1/2)\,\mathrm{Log}\,z}, D)$ and $(-e^{(1/2)\,\mathrm{Log}\,z}, D)$, respectively. Also, we observe that the domain formed by the union of all the domains D_t (which can be clearly covered by finitely many such disks), $0 \le t \le 2\pi$, surrounding the origin is not simply connected. In the case of a simply connected domain, the result of the continuation will be unique, no matter what chain is used. This is the substance of the Monodromy Theorem. ●

The difference between single-valued and multiple-valued functions may be viewed from another point of view. Suppose $f(z)$ is analytic in a domain D. A point z_1 is said to be a *regular point* of $f(z)$ if the function element (f, D) can be analytically continued along some curve from a point in D to the point z_1. The set of all regular points of $f(z)$ is called the *domain of regularity* for $f(z)$.

As we have seen, the function $f_0(z) = \sum_{n=0}^{\infty} z^n$ has domain of regularity $\{z : z \neq 1\}$. Note that the function $f(z) = 1/(1-z)$ is analytic in the domain of regularity for $f_0(z)$ and agrees with $f_0(z)$ at all points where they are both analytic.

Consider now the function

$$F_0(z) = \int_0^z f_0(\zeta)\, d\zeta = \int_0^z \left(\sum_{n=0}^{\infty} \zeta^n\right) d\zeta = \sum_{n=0}^{\infty} \frac{z^{n+1}}{n+1} \quad (|z| < 1),$$

where the path of integration lies in the unit disk. The function

$$F(z) = \int_0^z \frac{d\zeta}{1-\zeta} = -\,\mathrm{Log}\,(1-z)$$

agrees with $F_0(z)$ in the disk $|z| < 1$, and is analytic everywhere in the plane except $z = 1$ and the ray $\mathrm{Arg}\,(1-z) = \pi$ (i.e., the ray along the positive real axis beginning at $z = 1$). The function

$$F_1(z) = -\log(1-z) \qquad (0 < \arg(1-z) < 2\pi)$$

is a continuation of $F(z)$ from the half-plane $0 < \mathrm{Arg}\,(1-z) < \pi$ to the whole plane, excluding the point $z = 1$ and the ray $\mathrm{Arg}\,(1-z) = 0$.

Thus the domain of regularity for

$$F_0(z) = \sum_{n=0}^{\infty} \frac{z^{n+1}}{n+1}$$

is $\{z : z \neq 1\}$. Note, however, that there does not exist a function that is both analytic in the domain of regularity for $F_0(z)$ and agrees with $F_0(z)$ in the disk $|z| < 1$. As we shall see by the next theorem, this phenomenon occurs only because $\{z : z \neq 1\}$ is a multiply connected domain.

Remark 13.4. We say that the multiple-valued function $\log(1-z)$ is regular in the domain $\{z : z \neq 1\}$ because each such point is a regular point. Some authors allow multiple-valued functions to be analytic. Their definition of analytic then corresponds to our definition of regular. This next theorem shows us that a regular function is always single-valued (hence analytic) in a simply connected domain. ●

Theorem 13.5. (Monodromy Theorem) *Let D be a simply connected domain, and suppose $f_0(z)$ is analytic in a domain $D_0 \subset D$. If the function element (f_0, D_0) can be analytically continued along every curve in D, then there exists a single-valued function $f(z)$ that is analytic throughout D with $f(z) \equiv f_0(z)$ in D_0.*

Proof. We outline the proof, leaving some details for the interested reader. Suppose the conclusion is false. Then there exist points $z_0 \in D_0, z_1 \in D$, and curves C_1, C_2 both having initial point z_0 and terminal point z_1 such that (f_0, D_0) leads to a different function element in a neighborhood of z_1 when analytically continued along C_1 than when analytically continued along C_2 (see Figure 13.6). This means that (f_0, D_0) does not return to the same function element when analytically continued along the closed curve $C_1 - C_2$.

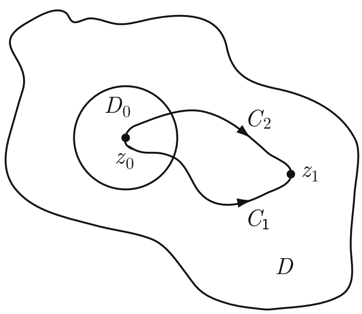


Figure 13.6.

To prove the theorem, it thus suffices to show that the function element (f_0, D_0), $D_0 \subset D$, can be continued along any closed curve lying in D and return to the same value. In the special case that the closed curve C is a rectangle, the proof will resemble that of Theorem 7.39.

Divide the rectangle C into four congruent rectangles, as illustrated in Figure 7.16. Continuation along C produces the same effect as continuation along these four rectangles taken together. If the conclusion is false for C, then it must be false for one of the four sub-rectangles, which we denote by C_1. We then divide C_1 into four congruent rectangles, for one of which the conclusion is false. Continuing the process, we obtain a nested sequence of rectangles for which the conclusion is false. According to Lemma 2.25, there is exactly one point, call it z^*, belonging to all the rectangles in the nest.

Since $z^* \in D$, there exists a function element (f^*, D^*) with $z^* \in D^* \subset D$. For n sufficiently large, the rectangle C_n of the nested sequence is contained in D^*. But this means that $f^*(z)$ is analytic in a domain containing C_n, contrary to the way C_n was defined. This contradiction concludes the proof in the special case in which the curve is a rectangle. For the general proof, see Hille [Hi]. ■

Suppose $f(z)$ is analytic in a domain D and z_0 is a boundary point of D. The point z_0 will be a regular point of $f(z)$ if, for some disk D_0 centered at z_0, there is a function element (f_0, D_0) such that $f_0(z) \equiv f(z)$ in the domain $D_0 \cap D$. Any boundary point of D that is not a regular point of $f(z)$ is said to be a *singular point* of $f(z)$.

For the function $f(z) = \sum_{n=0}^{\infty} z^n$ $(|z| < 1)$, we have seen that each point on the circle $|z| = 1$ is a regular point except for the point $z = 1$. That all points on the circle cannot be regular is a consequence of the following theorem.

Theorem 13.6. *If the radius of convergence of the series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is R, then $f(z)$ has at least one singular point on the circle $|z| = R$.*

Proof. Denote the disk $|z| < R$ by D, and suppose that all points on $|z| = R$ are regular points. Then, for each point z_α on the circle, we can find a function f_α defined in a disk D_α centered at z_α such that the function element (f_α, D_α) is a direct analytic continuation of (f, D). Since $\cup_\alpha D_\alpha$ covers the compact set $|z| = R$, a finite subcover $(D_1, D_2, \ldots, D_n)$ may be found. The function g defined by

$$g(z) = \begin{cases} f(z) & \text{if } z \in D \\ f_i(z) & \text{if } z \in D_i, \end{cases}$$

is analytic in the domain $D' = D \cup D_1 \cup D_2 \cup \cdots \cup D_n$. Since D' contains the disk $|z| \le R$, the domain must also contain the disk $|z| \le R + \epsilon$ for some positive ϵ. Hence the power series representation $g(z) = \sum_{n=0}^{\infty} a_n z^n$ is valid in the disk $|z| < R + \epsilon$, contradicting the fact that the Maclaurin series for $f(z)$ has radius of convergence R. ■

Corollary 13.7. *If $f(z)$ is analytic in the disk $|z - z_0| < R$ and the Taylor series expansion about $z = z_0$ has radius of convergence R, then $f(z)$ has at least one singular point on the circle $|z - z_0| = R$.*

Proof. Set $\zeta = z - z_0$, and apply the theorem to $f(\zeta)$. ■

Although we are guaranteed that a power series must have singular points on its circle of convergence, determining their location is, in general, a difficult problem. By placing a restriction on the coefficients, we can locate a particular singular point. Here is one of the results that we have in this direction.

Theorem 13.8. *Suppose* $f(z) = \sum_{n=0}^{\infty} a_n z^n$ *has radius of convergence* $R < \infty$*. If* $a_n \geq 0$ *for every* n*, then* $z = R$ *is a singular point of* f*.*

Proof. If $z = R$ is not a singular point, then $f(z)$ is analytic in some disk $D_0 : |z - R| < \epsilon$. For a positive number $\rho\,(< R)$ sufficiently close to R, we can find an open disk D_1 centered at $z = \rho$ that contains the point $z = R$ and is contained in D_0. Then the Taylor series

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(\rho)}{n!}(z-\rho)^n \tag{13.2}$$

converges at a point $z = R + \delta$ $(\delta > 0)$ (see Figure 13.7).

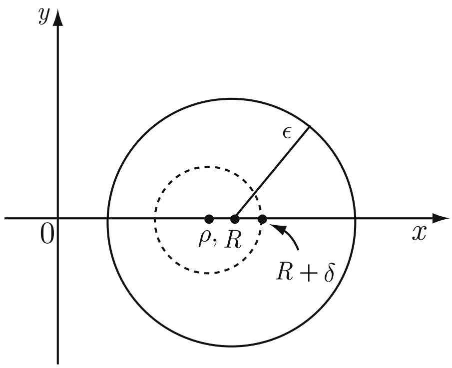


Figure 13.7.

According to Theorem 13.6, the series $\sum_{n=0}^{\infty} a_n z^n$ has a singular point somewhere on the circle $|z| = R$, say $Re^{i\theta_0}$. Hence the Taylor series

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(\rho e^{i\theta_0})}{n!}(z-\rho e^{i\theta_0})^n$$

has radius of convergence $R - \rho$ (if the radius of convergence were larger, then $Re^{i\theta_0}$ would not be a singular point). Note that for each n we have

$$f^{(n)}(\rho e^{i\theta_0}) = \sum_{k=n}^{\infty} k(k-1)\,\cdots\,(k-n+1)a_k(\rho e^{i\theta_0})^{k-n}. \tag{13.3}$$

Since $a_n \geq 0$, we obtain from (13.3) the inequality

$$\left| f^{(n)}(\rho e^{i\theta_0}) \right| \leq \sum_{k=n}^{\infty} k(k-1)\,\cdots\,(k-n+1)a_k\rho^{k-n} = f^{(n)}(\rho).$$

Thus

$$\frac{1}{R-\rho} = \limsup_{n\to\infty} \left| \frac{f^{(n)}(\rho e^{i\theta_0})}{n!} \right|^{1/n} \leq \limsup_{n\to\infty} \left| \frac{f^{(n)}(\rho)}{n!} \right|^{1/n},$$

which means that the radius of convergence of (13.2) is at most $R-\rho$. This contradicts the fact that the series converges at $z = R+\delta$. Therefore, $z = R$ is a singular point of $f(z)$. ∎

We have shown that a power series must have at least one singular point on its circle of convergence. The question arises as to whether there is an upper bound on the number of singular points on the circle. We will show that it is possible for every such point to be singular. If $f(z)$ is analytic in a domain whose boundary is C, and every point on C is a singular point of $f(z)$, then C is said to be the *natural boundary* of $f(z)$. In such a case, the domain of regularity is the same as the domain of analyticity.

We will make use of the following lemma in constructing a power series with a natural boundary.

Lemma 13.9. *Suppose that $f(z) = \sum_{n=0}^{\infty} a_n z^n$ has a radius of convergence R. If $f(re^{i\theta_0}) \to \infty$ as $r \to R$, then the point $Re^{i\theta_0}$ is a singular point of $f(z)$.*

Proof. If $Re^{i\theta_0}$ is a regular point, then there is a function $g(z)$ that is analytic in a disk centered at $Re^{i\theta_0}$ and agrees with $f(z)$ for $|z| < R$. But then

$$\lim_{r\to R^-} f(re^{i\theta_0}) = \lim_{r\to R^-} g(re^{i\theta_0}) = g(Re^{i\theta_0}),$$

contradicting the fact that the limit on the left side is infinite. ∎

Consider now the function

$$f(z) = \sum_{n=0}^{\infty} z^{2^n} = z + z^2 + z^4 + z^8 + \cdots,$$

which converges (and so is analytic) in the disk $|z| < 1$. We will show that the circle $|z| = 1$ is a natural boundary for the function $f(z)$. First observe that $f(z) \to \infty$ as $z \to 1$ along the real axis, so that $z = 1$ is a singular point (this is also a consequence of Theorem 13.8). Note that $f(z)$ satisfies the relation $f(z) = z + f(z^2)$. Hence $f(z)$ and $f(z^2)$ simultaneously approach ∞. But then $f(z^2) \to \infty$ when $z^2 \to 1$ through real values, thereby making -1 a singular point. This gives insight into the general method. The function $f(z)$ satisfies the recursive relationship

$$f(z) = z + z^2 + z^4 + \cdots + z^{2^{n-1}} + f(z^{2^n}).$$

For each fixed n, we have

$$|f(z)| \geq \left| f(z^{2^n}) \right| - n \quad (|z| < 1).$$

Since $f(z^{2^n}) \to \infty$ along each ray tending to a 2^n-th root of unity, it follows that each 2^n-th root of unity is a singular point. That is, all points of the form $e^{(2k\pi/2^n)i}$, where k and n are positive integers, are singular points. Now every neighborhood of any other point on the unit circle must contain one of these 2^n-th roots of unity. Hence no point on the unit circle is a regular point. That is, $|z| = 1$ is a natural boundary for $f(z)$.

A similar argument may be used for

$$f(z) = \sum_{n=0}^{\infty} z^{n!},$$

which is analytic in the disk $|z| < 1$. If $z = re^{2\pi(p/q)i}$, where p and q are positive integers and $0 < r < 1$, then (since $e^{2\pi(p/q)n!i} = 1$ for all $n \geq q$) it follows that

$$|f(z)| = \left| \sum_{n=0}^{q-1} r^{n!} e^{2\pi(p/q)n!i} + \sum_{n=q}^{\infty} r^{n!} \right| \geq \sum_{n=q}^{\infty} r^{n!} - q. \tag{13.4}$$

Since the right-hand side of (13.4) tends to ∞ as r tends to 1, all points of the form $e^{2\pi(p/q)i}$ are singular points. But these points are dense on $|z| = 1$, so that the unit circle is a natural boundary for $f(z)$.

Since a power series converges in a disk, its boundary must be a circle. But we have defined natural boundary to include a function for which the domain of analyticity need not be a disk. Consider the function

$$f(z) = \sum_{n=0}^{\infty} e^{-n!z}.$$

Since the series converges uniformly for $\operatorname{Re} z \geq \delta > 0$, the function $f(z)$ is analytic for $\operatorname{Re} z > 0$. We now show that the imaginary axis is a natural boundary for $f(z)$.

Suppose $z = x + 2\pi(p/q)i$, where p is an integer, q is a positive integer, and x is a positive real number. Then

$$|f(z)| = \left| \sum_{n=0}^{q-1} e^{-n!(x+2\pi(p/q)i)} + \sum_{n=q}^{\infty} e^{-n!x} \right| \geq \sum_{n=q}^{\infty} e^{-n!x} - q. \tag{13.5}$$

Because the right side of (13.5) tends to ∞ as x tends to 0, it follows that all points of the form $2\pi(p/q)i$ are singular points. But these points are dense on the imaginary axis so that the imaginary axis furnishes us with a natural boundary for $f(z)$.

Remark 13.10. Let Δ be the unit disk $|z| < 1$ and let $\gamma : [0, 1] \to \Delta$ be a curve with $\gamma(0) = 0$ and D be such that $0 \in D \subsetneq \Delta$. Then there is always an analytic continuation of $(\sum_{n=0}^{\infty} z^{n!}, \Delta)$ along γ. However, if $\gamma_1 : [0, 1] \to \mathbb{C}$

is given by $\gamma_1(t) = 2it$, there is no analytic continuation of $(\sum_{n=0}^{\infty} z^{n!}, \varDelta)$ along γ_1.

Similar comments apply for the function element $(\sum_{n=0}^{\infty} z^{2^n}, \varDelta)$. ●

Questions 13.11.

1. If $f(z) = z$ in a domain D_0, can $f(z)$ be analytic in a domain D_1 even though $f(z) \neq z$ in D_1?
2. Can two functions, analytic in the disk $|z| < 1$, agree at infinitely many points there and not agree everywhere in the disk?
3. Can an analytic continuation always be transformed into a direct analytic continuation?
4. Is it possible that the function elements (f, D_1) and (g, D_2) can be connected by an infinite chain of function elements, but by no finite subchain?
5. Why is the domain of regularity a domain?
6. What is the difference between a singular point and a singularity? A regular point and a point of analyticity?
7. Can infinitely many points on the boundary C of a domain be singular without C being a natural boundary?
8. If $D_1, D_2, \dots, D_n$ are domains, when is their union a domain?
9. Is the converse of Lemma 13.9 true?
10. Is there a relationship between gaps in the coefficients of the Maclaurin series for $f(z)$ and the circle of convergence being a natural boundary?
11. Is there a relationship between the Cauchy Theorem and the Monodromy Theorem?
12. What does the Monodromy theorem tell us about $\log z$? About $\sqrt{z}$?

Exercises 13.12.

1. Given a set of real numbers $0 \leq \theta_1 < \theta_2 < \cdots < \theta_n < 2\pi$, construct a function $f(z)$ such that
   (i) $f(z)$ is analytic in $|z| < 1$;
   (ii) the only singular points of $f(z)$ on the unit circle are at $e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_n}$.
2. Given (f_1, D_1), where $f_1(z) = \sum_{n=0}^{\infty} z^n$ and $D_1 = \{|z| < 1\}$, construct a chain $\{(f_1, D_1), (f_2, D_2), \dots, (f_n, D_n)\}$.
3. Show that the set of regular points of an analytic function is open, and the set of singular points is closed.
4. (a) Show that $f(z) = \sum_{n=0}^{\infty}[z^{2^{n+1}}/(1-z^{2^{n+1}})]$ is analytic in the domain $|z| < 1$ and the domain $|z| > 1$, and that $|z| = 1$ is a natural boundary for the function in each domain.
   (b) Determine $f(z)$ in each of these domains in closed form.
5. Show that $|z| = 1$ is a natural boundary for $\sum_{n=0}^{\infty} z^{3^n}$.
6. Suppose $\sum_{n=0}^{\infty} a_n z^{n!}$ $(a_n > 0)$ has radius of convergence R. Show that $|z| = R$ is a natural boundary.

7. Suppose $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is analytic for $|z| < 1$ and that a_n is real for each n. If $\sum_{n=1}^{k} a_n \to \infty$ as $k \to \infty$, show that $z = 1$ is a singular point for $f(z)$.
8. Suppose $f(z) = \sum_{n=0}^{\infty} a_n z^n$ has radius of convergence 1 and that the only singularities on the circle $|z| = 1$ are simple poles. Show that the sequence $\{a_n\}$ is bounded.
9. Show that $f(z) = \int_0^1 (1 - tz)^{-1}\, dt$ is an analytic continuation of $f_0(z) = \sum_{n=1}^{\infty} z^{n-1}/n$ from the unit disk $|z| < 1$ into the whole complex plane minus the interval $[1, \infty)$.
10. Suppose $f(z) = \sum_{n=0}^{\infty} (-1)^n a_n z^n$ has radius of convergence R and $a_n \geq 0$ for every n. Show that $z = -R$ is a singular point.

13.2 Special Functions

There are functions which arise so frequently in complex analysis that they have intrinsic interest. The gamma function of Euler and the zeta function of Riemann are two such “special functions” which require special attention. As we have seen in the previous chapter, the gamma function is meromorphic with simple poles at $0, -1, -2, \ldots$, and it is free of zeros. Its reciprocal is an entire function, with a simple zero at each nonpositive integers and with no other zeros. This may be expressed as

$$\frac{1}{\Gamma(z)} = z e^{\gamma z} \prod_{k=1}^{\infty} \left(1 + \frac{z}{k}\right) e^{-z/k}, \tag{13.6}$$

where

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \ln n \right).$$

Thus we may rewrite (13.6) as

$$\begin{aligned}
\frac{1}{\Gamma(z)} &= \left[\lim_{n\to\infty} z e^{[1+(1/2)+(1/3)+\cdots+(1/n)]z - z\ln n} \right] \lim_{n\to\infty} \prod_{k=1}^{n} \left(\frac{z+k}{k} \right) e^{-z/k} \\
&= \lim_{n\to\infty} \left[z e^{-z\ln n} \prod_{k=1}^{n} \left(1 + \frac{z}{k}\right) \right] \\
&= \lim_{n\to\infty} \frac{z(z+1)(z+2)\cdots(z+n)}{n^z n!}.
\end{aligned}$$

This leads to an alternate expression for the gamma function, namely

$$\Gamma(z) = \lim_{n\to\infty} \frac{n! n^z}{z(z+1)\cdots(z+n)}, \tag{13.7}$$

which is defined for all values except zero and the negative integers. Equation (13.7) is referred to as "Gauss's formula". Therefore, for all values of z with $z \neq 0, -1, -2, \dots$, we get that

$$\Gamma(z+1) = \lim_{n\to\infty} \frac{nz}{z+n+1} \left(\frac{n!n^z}{z(z+1)\cdots(z+n)} \right) = z\Gamma(z).$$

In this way, we obtain an alternate proof of the functional equation of the gamma function shown in the previous chapter. There is still one more method to obtain this equation as we shall see soon.

In real analysis, the gamma function is defined in terms of the improper integral

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt \quad (x>0). \tag{13.8}$$

Note that the integral (13.8) makes no sense when $x \leq 0$. Indeed, as $e^{-t} > e^{-1}$ for all $t \in (0,1)$, and for $0<\delta<1$

$$\int_\delta^1 t^{x-1}e^{-t}\,dt \geq \frac{1}{e}\int_\delta^1 t^{x-1}\,dt = \frac{1}{e}\left(\frac{1-\delta^x}{x}\right)$$

which approaches ∞ as $\delta \to 0^+$ for $x<0$. Thus, the improper integral (13.8) diverges for $x<0$. It is easy to see that it also diverges at $x=0$.

To see that the integral (13.8) converges for all positive x, we write

$$\Gamma(x) = \int_0^1 t^{x-1}e^{-t}\,dt + \int_1^\infty t^{x-1}e^{-t}\,dt = I_1 + I_2.$$

Since $e^{-t} \leq 1$ for $t \geq 0$, it follows that the integral (13.8) converges at $t=0$ because for each $\delta>0$,

$$\int_\delta^1 t^{x-1}e^{-t}\,dt \leq \int_\delta^1 t^{x-1}\,dt = \frac{1-\delta^x}{x} < \frac{1}{x}$$

so that $I_1 \leq 1/x$. For large t,

$$t^{x-1}e^{-t} \leq e^{t/2}e^{-t} = e^{-t/2}$$

so that the integral converges at ∞. In fact, since $\lim_{t\to\infty}(t^{x-1}/e^t) = 0$, the integrand of I_2 is also bounded so that

$$\int_N^\infty t^{x-1}e^{-t}\,dt \leq \int_N^\infty \frac{t^{x-1}}{t^{x+1}}\,dt = \frac{1}{N} \quad (N \geq N(x)).$$

Hence $\Gamma(x)$ is defined for all $x>0$. An integration by parts gives

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t}\,dt = -\frac{t^x}{e^t}\bigg|_0^\infty + x\int_0^\infty t^{x-1}e^{-t}\,dt = x\Gamma(x). \tag{13.9}$$

Note that (13.6) has been shown to satisfy $\Gamma(x+1) = x\Gamma(x)$ for complex values of x. From (13.9) and the fact

$$\Gamma(1) = \int_0^\infty e^{-t}\,dt = 1,$$

it follows that $\Gamma(n+1) = n!$ for all positive integers n.

Consider now the complex-valued function

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,dt. \tag{13.10}$$

For $z = x + iy, x > 0$, we have

$$\left|t^{z-1}\right| = \left|e^{(x-1)\,\mathrm{Log}\,t + iy\,\mathrm{Log}\,t}\right| = e^{(x-1)\,\mathrm{Log}\,t} = t^{x-1}.$$

Hence the integral (13.10) converges absolutely for $x > 0$, with

$$|\Gamma(z)| \le \int_0^\infty \left|t^{z-1}e^{-t}\right|\,dt = \Gamma(x)$$

so that (13.10) is well defined in the half-plane $\mathrm{Re}\,z > 0$. We wish to show that (13.10) has two important properties: first, it is analytic for $\mathrm{Re}\,z > 0$; second, it agrees with (13.7) for $\mathrm{Re}\,z > 0$. This will justify the apparently inexcusable notation in which the same letter is used for (13.10) and (13.7).

Let K be a compact subset of the half-plane $\mathrm{Re}\,z > 0$. For $z = x+iy \in K$, choose x_0, x_1 so that $0 < x_0 \le x \le x_1 < \infty$. Then, we have

$$|\Gamma(z)| \le \Gamma(x) \le \int_0^1 t^{x_0-1}e^{-t}\,dt + \int_1^\infty t^{x_1-1}e^{-t}\,dt < \Gamma(x_0) + \Gamma(x_1).$$

Thus $\Gamma(z)$ is bounded in the infinite strip

$$x_0 \le \mathrm{Re}\,z \le x_1. \tag{13.11}$$

For $n \ge 1$, we set

$$\Gamma_n(z) = \int_{1/n}^n t^{z-1}e^{-t}\,dt.$$

We will show that $\Gamma_n(z)$ is analytic for $\mathrm{Re}\,z > 0$, with

$$\Gamma_n'(z) = \int_{1/n}^n t^{z-1}e^{-t}\ln t\,dt.$$

To this end, we show that, on any strip of the form (13.11), the expression

$$\begin{aligned}\left|\frac{\Gamma_n(z+h) - \Gamma_n(z)}{h} - \int_{1/n}^n t^{z-1}e^{-t}\ln t\,dt\right| &= \left|\int_{1/n}^n t^{z-1}e^{-t}\left(\frac{t^h-1}{h} - \ln t\right)dt\right| \\ &\le \int_{1/n}^n t^{x-1}e^{-t}\left|\frac{t^h-1}{h} - \ln t\right|dt\end{aligned}$$

can be made arbitrarily small for $|h|$ sufficiently small. Using the mean-value theorem and the uniform continuity of $\ln t$ on the interval $[1/n, n]$, we can show that $(t^h - 1)/h$ converges uniformly to $\ln t$ for $1/n \le t \le n$. It thus follows when $|h| < \delta(\epsilon)$ that the last integral above is bounded above by

$$\epsilon \int_{1/n}^{n} t^{x-1} e^{-t}\, dt < \epsilon \Gamma(x) < \epsilon(\Gamma(x_0) + \Gamma(x_1)).$$

Hence $\Gamma_n(z)$ is analytic (for $x_0 < \operatorname{Re} z < x_1$), with

$$\Gamma_n'(z) = \int_{1/n}^{n} t^{z-1} e^{-t} \ln t\, dt.$$

But

$$\lim_{n\to\infty} \Gamma_n(z) = \Gamma(z)$$

for $x_0 \le \operatorname{Re} z \le x_1$. Since $\Gamma_n(z)$ is locally uniformly bounded in the right half-plane, Montel's theorem (Theorem 11.14) may be applied to show that $\Gamma(z)$ is analytic for $\operatorname{Re} z > 0$.

We now show that the integral definition (13.10) agrees with (13.7) for $x = \operatorname{Re} z > 0$. Set

$$\Gamma_n^*(x) = \int_0^n t^{x-1} \left(1 - \frac{t}{n}\right)^n dt \quad (x > 0,\ n \ge 1).$$

Integrating by parts, we obtain

$$\begin{aligned} \Gamma_n^*(x) &= \frac{t^x}{x}\left(1 - \frac{t}{n}\right)^n \Bigg|_0^n + \frac{1}{x}\int_0^n t^x \left(1 - \frac{t}{n}\right)^{n-1} dt \\ &= \frac{1}{x}\int_0^n t^x \left(1 - \frac{t}{n}\right)^{n-1} dt. \end{aligned}$$

Integrating by parts $n-1$ more times, we get

$$\begin{aligned} \Gamma_n^*(x) &= \frac{1}{x}\frac{n-1}{n(x+1)}\frac{n-2}{n(x+2)} \cdots \frac{1}{n(x+n-1)} \times \int_0^n t^{x+n-1}\, dt \\ &= \frac{(n-1)!n^{x+n}}{n^{n-1}x(x+1)\cdots(x+n)} \\ &= \frac{n!n^x}{x(x+1)\cdots(x+n)}. \end{aligned}$$

Thus for $x > 0$,

$$\lim_{n\to\infty} \Gamma_n^*(x) = \lim_{n\to\infty} \frac{n!n^x}{x(x+1)\cdots(x+n)}. \tag{13.12}$$

If we can now show that

$$\lim_{n\to\infty} \Gamma_n^*(x) = \int_0^\infty t^{x-1}e^{-t}\,dt$$

on the interval $[1, 2]$, it will then follow from the identity theorem that

$$\lim_{n\to\infty} \frac{n!n^z}{z(z+1)\cdots(z+n)} = \int_0^\infty t^{z-1}e^{-t}\,dt$$

in the largest domain containing the interval $[1, 2]$ in which both functions are analytic; that is, the representations (13.7) and (13.10) will have been shown to be equal in the right half-plane.

For $n > N$, we have

$$\Gamma_n^*(x) > \int_0^N t^{x-1}\left(1-\frac{t}{n}\right)^n dt \quad (1 \le x \le 2). \tag{13.13}$$

The sequence of polynomials $f_n(t) = (1 - t/n)^n$ converges uniformly to e^{-t} on any finite interval $[a, b]$. Furthermore,

$$f_n(t) \le f_{n+1}(t) \le e^{-t}$$

for n sufficiently large. Hence for each fixed x, the integrand of (13.13) (as a function of t) converges uniformly to $t^{x-1}e^{-t}$ on the interval $[0, N]$. Therefore,

$$\lim_{n\to\infty} \Gamma_n^*(x) \ge \lim_{n\to\infty} \int_0^N t^{x-1}\left(1-\frac{t}{n}\right)^n dt = \int_0^N t^{x-1}e^{-t}\,dt.$$

Since N is arbitrary, it follows that

$$\lim_{n\to\infty} \Gamma_n^*(x) \ge \int_0^\infty t^{x-1}e^{-t}\,dt. \tag{13.14}$$

But $\Gamma_n^*(x) \le \int_0^n t^{x-1}e^{-t}\,dt \le \int_0^\infty t^{x-1}e^{-t}\,dt$, so that

$$\lim_{n\to\infty} \Gamma_n^*(x) \le \int_0^\infty t^{x-1}e^{-t}\,dt. \tag{13.15}$$

Combining (13.14) and (13.15), we see that (13.10) agrees with (13.7) for $1 \le x \le 2$, and consequently they must agree in the right half-plane. Hence (13.7) (or (13.6)) may be viewed as a direct analytic continuation of the function

$$\int_0^\infty t^{z-1}e^{-t}\,dt$$

from the domain $\operatorname{Re} z > 0$ to $\mathbb{C}\setminus\{0, -1, -2, \ldots\}$.

Our next discussion concerns the function

$$\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}, \tag{13.16}$$

known as the *Riemann-zeta function.* (Here we use the traditional notation denoting the complex variable $s = \sigma + it$ rather than $z = x + iy$.) This is one of the most challenging and fascinating functions which has a natural link connecting the set of prime numbers with analytic number theory. We have already met this series at $s = 2$ and $s = 4$ with (p. 433 and Exercise 12.30(6))

$$\zeta(2) = \pi^2/6 \quad \text{and} \quad \zeta(4) = \pi^4/90.$$

Since $f_n(s) = n^{-s} = e^{-s\,\mathrm{Log}\,n}$ is an entire function and for $s = \sigma + it$,

$$|n^{-s}| = e^{-\sigma\,\mathrm{Log}\,n} = n^{-\sigma},$$

we see that the series (13.16) converges absolutely for $\mathrm{Re}\,s > 1$ and uniformly for $\mathrm{Re}\,s \geq \sigma_0 > 1$. Hence $\zeta(s)$ represents an analytic function in the half-plane $\mathrm{Re}\,s > 1$. Consequently,

$$\zeta'(s) = \sum_{n=1}^{\infty} f_n'(s) = -\sum_{n=2}^{\infty} (\ln n) n^{-s} \quad \text{for } \mathrm{Re}\,s > 1,$$

and more generally,

$$\zeta^{(k)}(s) = (-1)^k \sum_{n=2}^{\infty} (\ln n)^k n^{-s} \quad \text{for } \mathrm{Re}\,s > 1.$$

Now, to see its link with the collection of prime numbers, we prove the following

Theorem 13.13. (Euler's Product Formula) *For $\sigma > 1$, the infinite product $\prod_p (1 - p^{-s})$ converges and*

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right), \tag{13.17}$$

where the product is taken over the set $P = \{2, 3, 5, 7, 11, \dots\}$ of all prime numbers p.

Proof. Since the series $\sum p^{-s}$ converges absolutely for all $\mathrm{Re}\,s > 1$, and it converges uniformly on every compact subset of the half-plane $\mathrm{Re}\,s > 1$, the infinite product (13.17) converges. Next we note that for $\sigma > 1$

$$\zeta(s)\frac{1}{2^s} = \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + \cdots$$

so that

$$\zeta(s)\left(1 - \frac{1}{2^s}\right) = 1 + \frac{1}{3^s} + \frac{1}{5^s} + \cdots .$$

Similarly, one can find that

$$\zeta(s)\left(1-\frac{1}{2^s}\right)\left(1-\frac{1}{3^s}\right)=1+\frac{1}{5^s}+\frac{1}{7^s}+\frac{1}{11^s}+\cdots.$$

More generally,

$$\zeta(s)\left(1-2^{-s}\right)\left(1-3^{-s}\right)\cdots\left(1-p_N^{-s}\right)=\sum m^{-s}=1+p_{N+1}^{-s}+\cdots$$

and because of the unique factorization of integers, we can continue the procedure to obtain in the limiting case

$$\zeta(s)\prod_{p,\,\text{prime}}\left(1-p^{-s}\right)=1,$$

as desired. ■

Next, we wish to find an analytic continuation of the function element

$$(\zeta(s),\,\operatorname{Re}s>1).$$

To do this, we will first establish a connection between the Riemann-zeta function and the gamma function. Recall the integral representation

$$\Gamma(s)=\int_0^\infty x^{s-1}e^{-x}\,dx\quad(\operatorname{Re}s>0).$$

The substitution $x=nt$ gives

$$\Gamma(s)=n^s\int_0^\infty t^{s-1}e^{-nt}\,dt. \tag{13.18}$$

Applying the identity

$$\sum_{n=1}^{k}e^{-nt}=e^{-t}\left(\frac{1-e^{-kt}}{1-e^{-t}}\right)=\frac{1-e^{-kt}}{e^t-1}$$

to (13.18), we get

$$\sum_{n=1}^{k}\frac{1}{n^s}=\frac{1}{\Gamma(s)}\int_0^\infty\frac{1-e^{-kt}}{e^t-1}t^{s-1}\,dt.$$

Thus for $\operatorname{Re}s>1$ and k a positive integer, we have

$$\sum_{n=1}^{k}\frac{1}{n^s}=\frac{1}{\Gamma(s)}\int_0^\infty\frac{t^{s-1}}{e^t-1}\,dt-\frac{1}{\Gamma(s)}\int_0^\infty\frac{e^{-kt}}{e^t-1}t^{s-1}\,dt \tag{13.19}$$

because both integrals converge. It will now be shown that the last integral tends to 0 as $k\to\infty$.

Since $|t^{s-1}|=t^{\sigma-1}$ ($\sigma=\operatorname{Re}s$), it follows that

$$\left|\int_0^\infty \frac{e^{-kt}}{e^t - 1} t^{s-1}\, dt\right| \leq \int_0^\infty \frac{e^{-kt}}{e^t - 1} t^{\sigma-1}\, dt. \tag{13.20}$$

Given $\epsilon > 0$, choose δ small enough so that

$$\int_0^\delta \frac{e^{-kt}}{e^t - 1} t^{\sigma-1}\, dt \leq \int_0^\delta \frac{t^{\sigma-1}}{e^t - 1}\, dt < \epsilon \tag{13.21}$$

for all k. Next choose k large enough so that

$$\int_\delta^\infty \frac{e^{-kt}}{e^t - 1} t^{\sigma-1}\, dt \leq e^{-k\delta} \int_\delta^\infty \frac{t^{\sigma-1}}{e^t - 1}\, dt < \epsilon. \tag{13.22}$$

Combining (13.21) and (13.22), we see that the integral in (13.20) becomes arbitrarily small for k sufficiently large. Upon letting k approach ∞ in (13.19), we obtain the following result which relates the zeta function with the gamma function.

Theorem 13.14. *For* $\operatorname{Re} s > 1$,

$$\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s} = \frac{1}{\Gamma(s)} \int_0^\infty \frac{t^{s-1}}{e^t - 1}\, dt. \tag{13.23}$$

The problem of extending the domain of definition for the Riemann-zeta function is that the integral in (13.23) diverges for $\operatorname{Re} s \leq 1$. Now, we wish to extend $\zeta(s)$ analytically to be a meromorphic function on $\mathbb{C}$ with a simple pole at $s = 1$. To do this, we represent $\zeta(s)$ as a contour integral with the help of (13.23) that avoids the origin, so that the resultant function will be shown to be entire. The continuation will then be accomplished by relating this new function to the integral in (13.23).

Let C_ϵ consist of the part of the positive real axis from ∞ to ϵ $(0 < \epsilon < 2\pi)$, the circle centered at the origin of radius ϵ traversed in the counterclockwise direction, and the positive real axis from ϵ to ∞ (see Figure 13.8). Notice that the contour is chosen with the usual positive orientation. Compare the contour in Figure 13.8 with the contour in Figure 9.6. Write

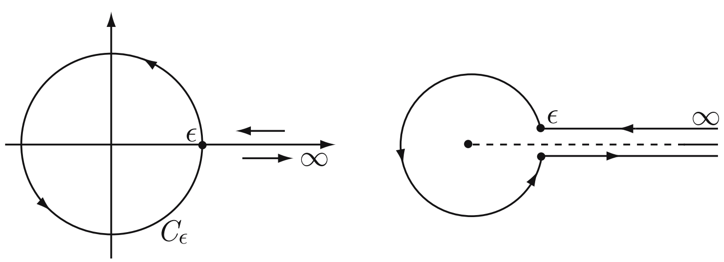


Figure 13.8.

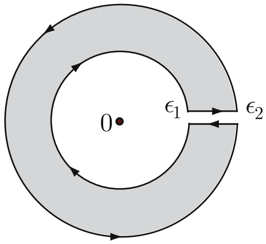


Figure 13.9.

$$z^{s-1} = e^{(s-1)\log z},$$

where $\log z = \ln|z| + i\arg z$ with $\arg z = 0$ when z lies in top edge of the branch cut of the real axis from ∞ to ϵ, whereas $\arg z = 2\pi$ when z lies on the bottom edge of branch cut from ϵ to ∞. Now, consider the function

$$f(s) = \int_{C_\epsilon} \frac{z^{s-1}}{e^z - 1}\, dz \quad (0 < \epsilon < 2\pi). \tag{13.24}$$

The integral converges, and it represents an entire function of s. Note that the value of the integral in (13.24) is independent of ϵ. To see this, suppose that $0 < \epsilon_1 < \epsilon_2 < 2\pi$. The region $C_{\epsilon_2} - C_{\epsilon_1}$, illustrated in Figure 13.9, is seen to be simply connected. Cauchy's theorem may thus be applied to show that

$$\int_{C_{\epsilon_2} - C_{\epsilon_1}} \frac{z^{s-1}}{e^z - 1}\, dz = \int_{C_{\epsilon_2}} \frac{z^{s-1}}{e^z - 1}\, dz - \int_{C_{\epsilon_1}} \frac{z^{s-1}}{e^z - 1}\, dz = 0.$$

Therefore,

$$\int_{C_{\epsilon_2}} \frac{z^{s-1}}{e^z - 1}\, dz = \int_{C_{\epsilon_1}} \frac{z^{s-1}}{e^z - 1}\, dz.$$

To evaluate the integral in (13.24), we first assume that $\operatorname{Re} s > 1$, and we express it in the form

$$\begin{aligned} f(s) &= \int_\infty^\epsilon \frac{t^{s-1}}{e^t - 1}\, dt + \int_{|z|=\epsilon} \frac{z^{s-1}}{e^z - 1}\, dz + \int_\epsilon^\infty \frac{e^{(s-1)(\ln t + 2\pi i)}}{e^t - 1}\, dt \quad (13.25) \\ &= (e^{2\pi i s} - 1) \int_\epsilon^\infty \frac{t^{s-1}}{e^t - 1}\, dt + \int_{|z|=\epsilon} \frac{z^{s-1}}{e^z - 1}\, dz. \end{aligned}$$

Suppose that $\sigma = \operatorname{Re} s > 1$. From the identity

$$e^z - 1 = z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

we see that for $|z|$ sufficiently small, $|e^z - 1| \geq |z|/2$. Hence

$$\left| \int_{|z|=\epsilon} \frac{z^{s-1}}{e^z - 1}\, dz \right| \le 2 \int_{|z|=\epsilon} \frac{\epsilon^{\sigma-1}}{\epsilon}\, |dz| = 4\pi\epsilon^{\sigma-1},$$

which approaches 0 as $\epsilon \to 0$. Hence, for $\operatorname{Re} s > 1$, $f(s)$ tends to a limit as $\epsilon \to 0$. Since $f(s)$ is independent of ϵ, we may evaluate $f(s)$ by letting $\epsilon \to 0$ in (13.25). This yields

$$f(s) = (e^{2\pi is} - 1) \int_0^\infty \frac{t^{s-1}}{e^t - 1}\, dt \quad (\operatorname{Re} s > 1). \tag{13.26}$$

We can compare (13.26) with (13.23) to get

Theorem 13.15. *For the branch of z^{s-1} and the contour C_ϵ indicated above, we have*

$$\zeta(s) = \frac{f(s)}{(e^{2\pi is} - 1)\Gamma(s)} \quad (\operatorname{Re} s > 1). \tag{13.27}$$

That is, we have an identity valid for all $\operatorname{Re} s > 1$*:*

$$\zeta(s) = \frac{1}{(e^{2\pi is} - 1)\Gamma(s)} \int_{C_\epsilon} \frac{z^{s-1}}{e^z - 1}\, dz.$$

Although (13.27) was proved only for $\operatorname{Re} s > 1$, as $f(s)$ is an entire function, the identity theorem may be used to extend this to a larger domain. Each simple pole of $\Gamma(s)$ is cancelled by a simple zero of $e^{2\pi is} - 1$. Hence $(e^{2\pi is} - 1)\Gamma(s)$ is an entire function. We have thus expressed $\zeta(s)$ in (13.27) as the quotient of entire functions, that is, as a meromorphic function. The poles of $\zeta(s)$ must occur at points where

$$(e^{2\pi is} - 1)\Gamma(s) = 0.$$

Now $\Gamma(s) \neq 0$, and $e^{2\pi is} - 1 = 0$ at the integers. But for zero and the negative integers, we have the zeros of $e^{2\pi is} - 1$ being cancelled by the poles of $\Gamma(s)$. Hence the only zeros of $(e^{2\pi is} - 1)\Gamma(s)$ occur at the positive integers. However, $\zeta(s)$ was already shown to be analytic for $\operatorname{Re} s > 1$ (thus, $f(s) = 0$ for $s = 2, 3, 4, \dots$). In conclusion, the ζ function is analytic for all values of s except $s = 1$ and hence, it continues analytically to $\mathbb{C} \setminus \{1\}$.

Therefore, the only possible pole for $\zeta(s)$ occurs at $s = 1$. To prove that $s = 1$ actually is a pole, we must show that $f(1) \neq 0$. From (13.24), we see that

$$f(1) = \int_{C_\epsilon} \frac{1}{e^z - 1}\, dz.$$

Since the only singularity of $1/(e^z - 1)$ inside C_ϵ $(0 < \epsilon < 2\pi)$ is a simple pole at $z = 0$, an application of the residue theorem shows that

$$f(1) = 2\pi i \lim_{z \to 0} \frac{z}{e^z - 1} = 2\pi i \neq 0.$$

Hence $\zeta(s)$ has a simple pole at $s = 1$ with residue

$$\lim_{s\to 1} \frac{(s-1)f(s)}{(e^{2\pi i s}-1)\Gamma(s)} = 2\pi i \lim_{s\to 1} \frac{s-1}{e^{2\pi i s}-1} = 1.$$

It follows that

$$\zeta(s) \sim \frac{1}{s-1} \quad \text{as } s \to 1\ .$$

Thus the equation (13.27) represents a direct analytic continuation of the series $\sum_{n=1}^{\infty} 1/n^s$ ($\mathrm{Re}\, s > 1$) to a function analytic in $\mathbb{C}$, except for a simple pole at $s = 1$. We have established the following

Theorem 13.16. *The zeta function is meromorphic in $\mathbb{C}$ with only simple pole at $s = 1$ with residue 1.*

It follows that the complete Riemann-zeta function may be expressed as

$$\zeta(s) = \frac{1}{s-1} + g(s),$$

where $g(s)$ is some entire function. Of course, for $\mathrm{Re}\, s > 1$,

$$g(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} - \frac{1}{s-1}.$$

In view of the identities

$$e^{2\pi i s} - 1 = 2ie^{\pi i s} \sin \pi s \quad \text{and} \quad \Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s},$$

(13.27) also takes the form

$$\zeta(s) = \frac{f(s)\Gamma(1-s)}{2\pi i e^{\pi i s}} = -\frac{\Gamma(1-s)}{2\pi i} e^{-\pi i (s-1)} f(s). \tag{13.28}$$

We may rewrite this as

$$\zeta(s) = -\frac{\Gamma(1-s)}{2\pi i} \int_{C_\epsilon} \frac{(-z)^{s-1}}{e^z - 1}\, dz \quad (0 < \epsilon < 2\pi),$$

where

$$(-z)^{s-1} = e^{(s-1)\,\mathrm{Log}\,(-z)} \quad \text{for } z \in \mathbb{C} \setminus [0, \infty).$$

The representations (13.27) and (13.28), though valid in $\mathbb{C}$, give no insight into the location of the zeros for the Riemann-zeta function. To aid us in this endeavor, we shall develop a recursive relationship for the Riemann-zeta function, providing explicit information, namely

Theorem 13.17. (Functional Equation of Zeta Function) *For all $s \in \mathbb{C}$, the ζ-function satisfies the functional equation*

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1-s)\zeta(1-s).$$

For the proof of this theorem, the following lemma will be helpful.

Lemma 13.18. *Let D be the domain consisting of the whole plane, excluding disks of the form $|z-2k\pi i| < \frac{1}{2}$, k an integer. Then there exists a real number $\delta > 0$ such that $|e^z - 1| \geq \delta$ for z in D.*

Proof. Since $e^z - 1$ is a periodic function of period $2\pi i$, it suffices to prove the inequality for the region R consisting of the strip $-\pi \leq \operatorname{Im} z \leq \pi$, excluding the disk $|z| < \frac{1}{2}$. Observe that $e^z - 1$ tends to ∞ as z approaches ∞ in the right half-plane on R, and approaches -1 as z tends to ∞ in the left half-plane of R. Choose δ, $0 < \delta < 1$, such that $|e^z - 1| \geq \delta$ on the circle $|z| = \frac{1}{2}$. Since $e^z - 1$ never vanishes in R, the minimum modulus theorem may be applied to show that $|e^z - 1| \geq \delta$ for all z in R. ■

Now we proceed to prove Theorem 13.17. For the proof, we modify the contour C_ϵ used to define $f(s)$ in (13.24). Fix s with $s < 0$. Let $0 < \epsilon < 2\pi$ and k be a positive integer. Let C_k differ from the contour in Figure 13.8 only in that the circle has radius $(2k+1)\pi$ instead of ϵ. Define

$$f_n(s) = \frac{1}{2\pi i} \int_{C_k} \frac{z^{s-1}}{e^z - 1}\, dz.$$

The idea is to relate the integral (13.24) defined for C_ϵ with a new integral defined for C_k but with a factor $1/2\pi i$, introduced for convenience. Then the function

$$\frac{z^{s-1}}{e^z - 1}$$

has simple poles inside the contour $C_k - C_\epsilon$ at the points

$$z = \pm 2n\pi i \quad (n = 1, 2, \ldots, k).$$

The residue at $\pm 2n\pi i$ is

$$\begin{aligned} \lim_{z \to \pm 2n\pi i} \left(\frac{z \pm 2n\pi i}{e^z - 1} \right) z^{s-1} &= (\pm 2n\pi i)^{s-1} \\ &= e^{(s-1)[\ln(2n\pi) + i \arg(\pm \pi i)]} \\ &= (2\pi)^{s-1} n^{s-1} e^{i(s-1)\arg(\pm \pi i)}. \end{aligned}$$

As $\arg(i\pi) = i\pi/2$ and $\arg(-i\pi) = i3\pi/2$, making use of the residue theorem, we have

$$\frac{1}{2\pi i} \int_{C_k - C_\epsilon} \frac{z^{s-1}}{e^z - 1}\, dz = (2\pi)^{s-1} \left(e^{(s-1)i\pi/2} + e^{(s-1)i3\pi/2} \right) \sum_{n=1}^{k} n^{s-1}.$$

We substitute

$$\begin{aligned} e^{(s-1)i\pi/2} + e^{(s-1)i3\pi/2} &= e^{(s-1)i\pi}\left(e^{-(s-1)i\pi/2} + e^{(s-1)i\pi/2}\right) \\ &= -2e^{i\pi s}\sin(\pi s/2). \end{aligned}$$

Hence it follows that

$$\frac{1}{2\pi i}\int_{C_k - C_\epsilon} \frac{z^{s-1}}{e^z - 1}\, dz = -2(2\pi)^{s-1} e^{i\pi s}\sin(\pi s/2)\sum_{n=1}^{k}\frac{1}{n^{1-s}}. \tag{13.29}$$

By Lemma 13.18, we have

$$\begin{aligned} \left|\frac{1}{2\pi i}\int_{|z|=(2k+1)\pi}\frac{z^{s-1}}{e^z - 1}\, dz\right| &\leq \frac{1}{2\pi\delta}\int_{|z|=(2k+1)\pi} |z^{s-1}|\,|dz| \\ &\leq \frac{1}{\delta}\{(2k+1)\pi\}^s \\ &\to 0 \quad \text{as } k\to\infty \text{ (since } s<0). \end{aligned} \tag{13.30}$$

In view of (13.30), we let $k \to \infty$ in (13.29) to obtain

$$\frac{1}{2\pi i}\int_{-C_\epsilon}\frac{z^{s-1}}{e^z - 1}\, dz = -2(2\pi)^{s-1} e^{i\pi s}\sin(\pi s/2)\zeta(1-s) \quad (s<0).$$

That is the function $f(s)$ in (13.24) may be expressed as

$$\frac{f(s)}{2\pi i} = 2(2\pi)^{s-1} e^{i\pi s}\sin(\pi s/2)\zeta(1-s) \quad (s<0). \tag{13.31}$$

A substitution of (13.31) into (13.28) yields the identity

$$\zeta(s) = 2^s \pi^{s-1}\sin\left(\frac{\pi s}{2}\right)\Gamma(1-s)\zeta(1-s) \quad (s<0). \tag{13.32}$$

Since both sides of this identity are meromorphic functions of s, this hold for all s, by the identity theorem. The proof of Theorem 13.17 is now complete.

Much of the analytic interest in the zeta function follows from the functional equation (13.32). For instance, the expression (13.32) enables us to locate some of the zeros of $\zeta(s)$. Note first that, as a consequence of Theorem 13.13, $\zeta(s)$ has no zeros in the half-plane $\mathrm{Re}\, s > 1$. Since $\Gamma(1-s)$ and $\zeta(1-s)$ are both analytic and nonzero for $\mathrm{Re}\, s < 0$, the only zeros of $\zeta(s)$ there are due to the zeros of $\sin(\pi s/2)$, that is, at the points $s = -2, -4, \ldots$. These zeros are called the *trivial zeros* of the Riemann-zeta function.

The only zeros unaccounted for must lie in the strip $0 \leq \mathrm{Re}\, s \leq 1$, which is called *the critical strip*. Now, we formulate

Theorem 13.19. *The only zeros of the zeta function not in the critical strip are at* $-2n$*,* $n \in \mathbb{N}$*.*

The problem of classifying the zeros of the zeta function is a formidable (and unsolved) task. Thus far, infinitely many zeros have been found in this strip; remarkably, all are situated on the line $\mathrm{Re}\, s = 1/2$, which is called *the critical line*. Also, it is proved that the zeta function has zeros neither on the line $\mathrm{Re}\, s = 1$ nor on the line $\mathrm{Re}\, s = 0$. A theorem of Hardy proves that there are infinitely many zeros inside the critical strip. See the books by Edwards [E] and Ivic [I] for further information.

Encouraging the student to further pursue mathematics, we end this book not with a theorem but with a famous conjecture known as the

Riemann Hypothesis. *All the nontrivial zeros of* $\zeta(s)$ *lie on the line* $\mathrm{Re}\, s = 1/2$.

Questions 13.20.

1. Is $\Gamma(0+) = \infty$?
2. Is $\int_{|z|=1/3} \Gamma(z)\, dz = 2\pi i$?
3. For $n \in \mathbb{N}$, what is the value of $\int_{|z|=n+1/3} \Gamma(z)\, dz$?
4. For $n \in \mathbb{N}$, is the function
$$\Gamma(z) - \frac{(-1)^n}{n!(z+n)}$$
analytic in the disk $|z+n| < 1$?
5. Is there a function $f(z) \neq \Gamma(z)$, analytic in the right half-plane, that satisfies the relationship $f(z+1) = zf(z)$?
6. What properties of the gamma function can most easily be proved by (13.7)?
7. What identities can be found by comparing (13.6), (13.7), and (13.10)?
8. In showing the equivalence of (13.7) and (13.10) in the right half-plane, why was it necessary to first show that they agreed on a finite interval?
9. What properties do the gamma function and the Riemann-zeta function have in common?
10. How do the properties of $\sum_{n=1}^{\infty} a_n z^n$ and $\sum_{n=1}^{\infty} (a_n/z^n)$ compare?
11. What kind of function is $(1-s)\zeta(s)/\Gamma(s)$?
12. What information about the Riemann-zeta function, other than the location of some zeros, can we obtain from (13.32)?
13. Is the set of all zeros of the zeta function symmetric with respect to both the critical line and the real axis?

Exercises 13.21.

1. Prove Legendre's duplication formula
$$\sqrt{\pi}\Gamma(2z) = 2^{2z-1}\Gamma(z)\Gamma(z+1/2).$$

2. Show that the gamma function may be expressed as

$$\Gamma(z) = \int_1^\infty e^{-t} t^{z-1}\, dt + \sum_{n=0}^\infty \frac{(-1)^n}{n!(z+n)}.$$

3. Show that $\operatorname{Re}\zeta(s) > 0$ when $\operatorname{Re} s \geq 2$.
4. Show that $(1-1/2^{s-1})\zeta(s)$ is an entire function and may be represented as $\sum_{n=1}^\infty (-1)^{n+1}/n^s$ for $\operatorname{Re} s > 1$. Where else does this series converge?
5. For $0 < \operatorname{Re} s < 1$, show that

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \left(\frac{1}{e^t - 1} - \frac{1}{t} \right) dt.$$

6. Show that $\zeta(1-s) = (1/2^{s-1}\pi^s) \cos(\pi s/2) \Gamma(s) \zeta(s)$.
7. Determine an analytic continuation of $\sum_{n=1}^\infty z^n/n^{1/4}$.
8. Consider the analytic function

$$f(z) = \sum_{n=1}^\infty \frac{1+c}{n+c} z^n \quad (c > -1).$$

Determine the largest domain to which f can be analytically continued? Determine an analytic continuation of f from the unit disk to a larger domain?