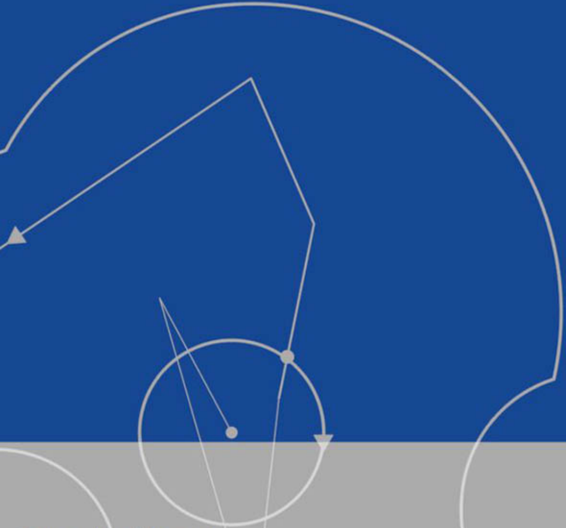
H. A. Priestley



Introduction to
# Complex Analysis

*second edition*

# Introduction to Complex Analysis

*This page intentionally left blank*

# Introduction to Complex Analysis

Second Edition

H. A. PRIESTLEY

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Taipei Toronto Shanghai

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan South Korea Poland Portugal
Singapore Switzerland Thailand Turkey Ukraine Vietnam

Published in the United States
by Oxford University Press Inc., New York



First published 1985
Revised edition 1990
Second edition 2003
Reprinted 2003, 2004, 2005 (twice)



A catalogue record for this title is available from the British Library

Library of Congress Cataloging in Publication Data

(Data available)

ISBN 0 19 852562 1 (Pbk) 978 0 19 852562 2
ISBN 0 19 852561 3 (Hbk) 978 0 19 852561 5

10 9 8 7 6 5

Printed in Great Britain
on acid-free paper by
Biddles Ltd,
King's Lynn, Norfolk

# Preface to the second edition

I have been gratified by the success achieved by *Introduction to complex analysis* since it first appeared in 1985, and have been pleasantly surprised by the range of users it has attracted. But any textbook shows its age in time and, nearly twenty years after its first publication, this one was certainly due for a make-over. The new edition is substantially different. from the revised edition published in 1990. I believe that the changes will enhance the book's suitability for a present-day readership. My overall aims, however, remain unchanged: 'to provide a text for a first course in complex analysis which is practical without being purely utilitarian and rigorous without being over-sophisticated or fussy.'

The new edition is considerably longer than the previous one, with additional discussion of key issues and extra examples and exercises. Extensive teaching experience has pinpointed for me where expanded or simplified explanations would particularly assist students. There is now a more protracted development of the early material, to take account of the likely knowledge base and mathematical experience of potential readers. Certain topics are treated in greater detail, to give students a thorough grounding in techniques that are used repeatedly; examples are Taylor series related to binomial expansions and zeros of holomorphic functions. As before, some familiarity with $\varepsilon$-$\delta$ real analysis is assumed. I have now taken *Real analysis*, by R. G. Bartle and D. R. Sherbert [**3**] as the core reference for this, but other texts would serve perfectly well.

Complex analysis is unusual amongst areas of mathematics in the range of mathematicians (and others) it attracts. It is intended that this book should be usable at several different levels and so serve a variety of readerships. The second edition has been structured to facilitate this. It is subdivided into very short chapters and much of the technical material has been positioned so that it can without loss be treated as optional. In addition, certain chapters and sections of chapters are designated 'basic track' and some as 'advanced track' (superseding the 'Level I' and 'Level II' designations in the original edition). So, for Cauchy's theorem and related results, Chapter 11 presents a basic track treatment adequate for all the applications, while the optional advanced track presentation in Chapter 12 explores the underlying ideas in greater depth.

The material has been re-arranged in such a way that the order in which topics can be studied is constrained as little as possible. Thus, at the extremes, it is feasible, for example,

- to take a geodesic route to Cauchy's theorem and its consequences, or
- to place emphasis in the early stages on geometric thinking, through a study

of mappings of the complex plane, or
- to concentrate on techniques rather than theory throughout, with a view primarily to developing skills required for applications.

The problem sets have been extensively revised and enlarged, and are carefully graded. Most of the exercises are quite elementary, and are designed to familiarize students with new concepts. Intermediate steps have been included in more challenging problems where experience has indicated the need for these. A few exercises introduce more advanced ideas and results.

Classic complex analysis is very much a triumph of 19th century mathematics. Nevertheless, there has quite recently been important research activity in, or related to, the subject. The famous Bieberbach conjecture was solved in 1984. Thanks to increasing computer power, the beauty of fractal curves has been revealed. Neither of these topics is suitable for inclusion in the core of the text. Instead, a brief appendix hints at these developments.

My thanks are due to those of my former students whose occasional blank looks have led me to work harder to explain certain points. I am grateful to the present undergraduates who have consumer-tested drafts of the new edition; here a special commendation goes to Ben Craig, who drew my attention to many more missing brackets, incorrect signs, and other small bugs than any of his peers. I am also grateful to a number of colleagues for constructive comments. I have, in part, been swayed by their suggestions that I should mention some notions excluded, perhaps too ruthlessly, from the first edition. But I remain unrepentant about the omission of topics I regard as too advanced for a first course. Finally, I should also like to thank the staff of Oxford University Press for their encouragement and support.

H. A. P.

*Oxford*
April 2003

# Preface to the first edition

This is a textbook for a first course in complex analysis. It aims to be practical without being purely utilitarian and to be rigorous without being oversophisticated or fussy. The power and significance of Cauchy's theorem—the centrepiece of complex analysis—is, I believe, best revealed initially through its applications. Consequently, emphasis has been put on those parts of the subject emanating from Cauchy's integral formula and Cauchy's residue theorem. This does not mean that the geometrical and topological aspects of complex analysis have been neglected, merely that it is recognized that a full appreciation of such concepts as index only comes with experience. Thus the chapters in which these important foundations are discussed are written in such a way that the student may at a first reading easily extract what he needs to proceed to the applications. He is, naturally, encouraged to return later in search of a deeper understanding.

The book is a metamorphosis of a set of notes in the series produced by the Mathematical Institute of the University of Oxford. As student opinion dictated it should, it betrays its previous incarnation—notably in its brevity and its style. Essential ideas are not submerged in a welter of details, material is locally arranged for ease of reference, and by-ways (however fascinating) are left unexplored. Advanced and specialized topics have been ruthlessly excluded. So, for example, analytic continuation and special functions receive only passing mention; a satisfactory treatment of either would have made unacceptable demands on the reader. Applied mathematicians have been provided with a thorough acccount of applicable complex analysis, but specific physical problems are not discussed. A chapter on Fourier and Lpalce transforms has been included. This is used to show off the techniques of residue calculus developed in the preceding chapters. It is also designed as a self-contained introduction to transform methods (and so strays somewhat beyond the confines of complex analysis), but does not purport to be a comprehensive survey of transform theory.

Some prior acquaintance with complex numbers is assumed. Apart from this, the only prerequisite is a course in elementary real analysis involving some exposure to $\varepsilon$-$\delta$ proofs. Many analysis and calculus texts cover the required background. I have taken K.G. Binmore, *Mathematical analysis: a straightforward introduction* [4] as my basic reference since it has the merit of having the same philosphy as the present book. Those concepts in real analysis which transfer, *mutatis mutandis*, from real analysis (continuity, etc.) are treated very briefly. Few students welcome, or benefit from, a detailed presentation of essentially familiar technical material. Also, the more time spent in these foothills, the less time is available for exploring the novel and spectacular terrain surrounding Cauchy's theorem.

Not all students will have the same mathematical background. To allow for this, I have adopted the convention that text enclosed in square brackets should be heeded by anyone to whom it makes sense but can safely be ignored by others. These occasional bracketed comments contain, for example, certain results in topology. It is accepted practice for texts on complex analysis to work with the Riemann integral rather than the Lebesgue integral. This is irritating for those who have graduated to the latter and confusing for those (Oxford students in particular) who are never taught the former. A dual approach is adopted here.To understand the book the reader needs a rudimentary knowledge of either Riemann integration or Lebesgue integration; signposts are provided for the followers of each theory.

Certain theorems have been designated with the customary proper names, but I have otherwise made no attempt to attribute theorems or proofs. Also, the subject has been so well worked over that I do not claim any originality for methods, examples, or exercises I happen never to have seen elsewhere. Among the books I have found most influential have been those by W. Rudin [**19**] and A.F. Beardon [**10**].

My preliminary notes on complex analysis evolved over about ten years. The first version fo these was based on some notes produced by Dr Ida Busbridge. She had earlier introduced me to 'complex variable', and I gratefully acknowledge my debt to her. It is also a pleasure to thank those colleagues in Oxford and elsewhere who directly or indirectly have had an influence on the book. However, my special thanks go to Dr Christine Farmer of London University; she has been involved with this project since its inception and has read draft after draft with care and patience. Her constructive criticisms have ben invaluable and her pencilled question marks unerring. Finally, thanks are due to Professor Michael Adams and Professor Michael Albert for their help with proof-reading, and to the staff of the Oxford University Press for encouraging me to write the book and for their assistance during its production.

H.A.P.

*Oxford*
March 1985

# Contents

# Notation and terminology

We use $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ to denote, respectively, the set of natural numbers, integers, real numbers, and complex numbers. Standard terms and symbols relating to sets and mappings have their conventional meanings. The following notation, which may not be universally familiar, is also used. Given sets $A$ and $B$, the set $\{a \in A : a \notin B\}$ is denoted $A \setminus B$, and given a mapping $f: A \to B$, we write the image set $\{f(a) : a \in A\}$ as $f(A)$. In addition, the **characteristic function**, $\chi_B$, of $B$ is given by $\chi_B(x) = 1$ if $x \in B$ and $\chi_B(x) = 0$ otherwise.

When we extend such concepts as differentiability from the real to the complex setting, we shall sometimes transfer secondary vocabulary and notation without comment. For example, once $f'(z)$ has been defined, we credit the reader with enough common sense to deduce what is meant by $f''(z)$ and $f^{(n)}(z)$.

The symbol $:=$ denotes 'equals by definition'; it is used to stress that an equation is defining something and also as a convenient shorthand. We denote the end of a proof by the customary symbol, $\square$. We adopt the Bourbaki dangerous bend symbol, ☡ to warn of a common pitfall. Finally, some of the more calculational sections of the book contain 'tactical tips', flagged by the symbol $\odot$. These explain various important points of strategy.

As explained in the preface, any comment in the text enclosed in square brackets is aimed just at those readers who have the knowledge to understand it.

*This page intentionally left blank*

# 1 The complex plane

Complex analysis has its roots in the algebraic, geometric, and topological structure of the complex plane. This chapter starts to explore these foundations. It is assumed that the reader has previously been introduced to complex numbers, and has had some practice in manipulating them. Consequently the first part of the chapter is designed to be a refresher course. It contains a summary of basic properties, presented without undue formality.

## Complex numbers

**1.1 Complex numbers.** A complex number is specified by a pair of real numbers $x$ and $y$; we write $x + \mathrm{i}y$, where i (sometimes also known as j) is a fixed symbol. (The arithmetical rules given in 1.4 force $\mathrm{i}^2 = -1$.) The set of complex numbers is denoted by $\mathbb{C}$. We use the customary abbreviations: $x$ for $x + \mathrm{i}0$, $\mathrm{i}y$ (or $y\mathrm{i}$) for $0 + \mathrm{i}y$, $0$ for $0 + \mathrm{i}0$, and i for $0 + \mathrm{i}1$. The first of these implies that we may regard $\mathbb{R}$, the real numbers, as a subset of $\mathbb{C}$. The terminology here is a legacy from the past: complex numbers are not complex, nor imaginary numbers any more imaginary than real numbers.

Two elements $x + \mathrm{i}y$ and $u + \mathrm{i}v$ of $\mathbb{C}$ are, by definition, equal if and only if $x = u$ and $y = v$. This allows us, given $z = x + \mathrm{i}y \in \mathbb{C}$, unambiguously to define $x$ to be the **real part** of $z$, written $\operatorname{Re} z$, and $y$ to be the **imaginary part** of $z$, written $\operatorname{Im} z$.

**1.2 Cartesian and polar representations.** It is convenient to represent complex numbers geometrically as points of a plane (the **complex plane**), also known as the **Argand diagram.** We equip the plane $\mathbb{R}^2$ in the usual way with Cartesian coordinate axes and identify $z = x + \mathrm{i}y$ with $(x, y) \in \mathbb{R}^2$. This is the **Cartesian representation.**

Alternatively, we may use polar coordinates and, for $(x, y) \in \mathbb{R}^2$, may write $x = r\cos\theta$ and $y = r\sin\theta$, where $r \geqslant 0$ and $\theta \in \mathbb{R}$. See Fig. 1.1. We write $\mathrm{e}^{\mathrm{i}\theta}$ as shorthand for $\cos\theta + \mathrm{i}\sin\theta$. Later, when we have introduced the complex

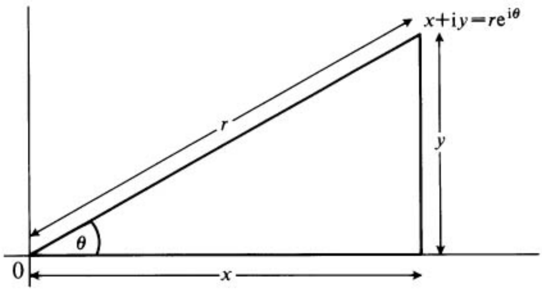



**Figure 1.1** Cartesian and polar representations

exponential function, we shall see that this is indeed $e^w$ evaluated at the point $w = i\theta$.

**1.3 Modulus and argument.** The **modulus** $|z|$ of $z = x + iy$ is defined to be

$$|z| = \sqrt{x^2 + y^2}$$

(where the positive square root is taken). This can be interpreted as the distance of $z$ from the origin, 0. Observe that

$$z = 0 \iff \operatorname{Re} z = \operatorname{Im} z = 0 \iff |z| = 0.$$

Because $\cos^2\theta + \sin^2\theta = 1$, we have $\left|e^{i\theta}\right| = 1$ for all real $\theta$. Now consider any $z \in \mathbb{C}$. Writing $z = x + iy$ in its polar form $re^{i\theta}$, we have $r = |z|$. For $z = 0$, we can choose $\theta$ arbitrarily. For $z \neq 0$, the angle $\theta$ is not unique: because of the periodicity of the functions cosine and sine, $\theta$ is only determined up to an integer multiple of $2\pi$. We call any value of $\theta$ with $z = re^{i\theta}$ **an argument** of $z = x + iy$. For now, we write $\arg z$ to denote any allowable value of $\theta$. Later we shall have to treat argument with much greater care. Indeed, the non-uniqueness of $\theta$, which may appear here merely as an inconvenience, turns out to have far-reaching consequences (see 7.10, 7.13, 10.4, 12.8–12.11).

Note especially the following very important and very useful facts: for $\theta \in \mathbb{R}$,

$$\begin{aligned}
e^{i0} &= 1 &\quad\text{and}\quad& e^{i\theta} = 1 \iff \theta = 2k\pi \quad (k \in \mathbb{Z});\\
e^{i\pi} &= -1 &\quad\text{and}\quad& e^{i\theta} = -1 \iff \theta = (2k+1)\pi \quad (k \in \mathbb{Z});\\
e^{i\pi/2} &= i &\quad\text{and}\quad& e^{i\theta} = i \iff \theta = \tfrac{1}{2}(2k+1)\pi \quad (k \in \mathbb{Z}).
\end{aligned}$$

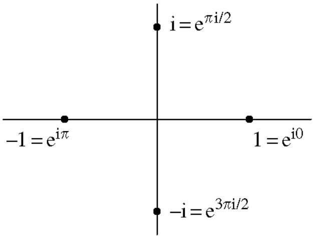



**Figure 1.2** Some special complex numbers

See Fig. 1.2. To prove these results we appeal to well-known facts about real trigonometric functions. For example, to obtain the first one we note that $\cos\theta = 1$ and $\sin\theta = 0$ both hold if and only if $\theta = 2k\pi$ $(k \in \mathbb{Z})$.

## Algebra in the complex plane

**1.4 The algebraic structure of the complex plane.** By extension of the corresponding operations for real numbers, addition and multiplication are defined in $\mathbb{C}$ by

$$(x + \mathrm{i}y) + (u + \mathrm{i}v) := (x + u) + \mathrm{i}(y + v),$$
$$(x + \mathrm{i}y)(u + \mathrm{i}v) := (xu - yv) + \mathrm{i}(xv + yu).$$

Taking $x = u = 0$ and $y = v = 1$, we obtain the identity $\mathrm{i}^2 = -1$.

Routine checking shows that the same arithmetical rules apply in $\mathbb{C}$ as in $\mathbb{R}$. For $z_1, z_2, z_3 \in \mathbb{C}$, we have commutative laws,

$$z_1 + z_2 = z_2 + z_1 \quad \text{and} \quad z_1 z_2 = z_2 z_1,$$

associative laws,

$$z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3 \quad \text{and} \quad z_1(z_2 z_3) = (z_1 z_2) z_3,$$

and the distributive law

$$z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3.$$

As expected, we have $0 + z = z$ and $1z = z$ for all $z \in \mathbb{C}$. In addition, given $z = x + \mathrm{i}y$, there exists

$$-z := (-x) + \mathrm{i}(-y)$$

such that $z+(-z) = 0$ and, so long as $z \neq 0$, there exists $1/z$ (also denoted $z^{-1}$) given by

$$1/z := x(x^2 + y^2)^{-1} - \mathrm{i}y(x^2 + y^2)^{-1},$$

which is such that $z(1/z) = 1$. We deduce from this a fact that we shall use frequently: $zw = 0$ implies $z = 0$ or $w = 0$ (for $z, w \in \mathbb{C}$).

In a mathematical nutshell, $\mathbb{C}$ forms a field. Informally, this simply means that the algebraic manipulation of complex numbers is just like that of real numbers, with the law $\mathrm{i}^2 = -1$ being available to simplify expressions.

**1.5 Products and powers; de Moivre's theorem.** It is worth noting that, while addition is most conveniently expressed using the Cartesian representation, the neatest formula for multiplication is in terms of the polar representation. To see this, take $z = r\mathrm{e}^{\mathrm{i}\theta}$ and $w = R\mathrm{e}^{\mathrm{i}\varphi}$. Write $z = r\cos\theta + \mathrm{i}r\sin\theta$ and $w = R\cos\varphi + \mathrm{i}R\sin\varphi$. Then, using the definition in 1.4, we have

$$\begin{aligned} zw &= rR(\cos\theta\cos\varphi - \sin\theta\sin\varphi) + \mathrm{i}rR(\cos\theta\sin\varphi + \cos\varphi\sin\theta) \\ &= rR\cos(\theta + \varphi) + \mathrm{i}rR\sin(\theta + \varphi), \end{aligned}$$

by standard trigonometric formulae. Hence $zw = rR\mathrm{e}^{\mathrm{i}(\theta+\varphi)}$. This implies in particular that $|zw| = |z|\,|w|$.

Let $z = r\mathrm{e}^{\mathrm{i}\theta}$ and let $n$ be a natural number. Then, from above and a routine proof by induction, we obtain

$$z^n = r^n\mathrm{e}^{\mathrm{i}n\theta}.$$

The special case $r = 1$ gives **de Moivre's theorem**:

$$(\cos\theta + \mathrm{i}\sin\theta)^n = \cos n\theta + \mathrm{i}\sin n\theta \quad (\theta \in \mathbb{R}).$$

Also, for $0 \neq z = r\mathrm{e}^{\mathrm{i}\theta}$, we have $1/z = r^{-1}\mathrm{e}^{-\mathrm{i}\theta}$ (write $1/z$ as $w = R\mathrm{e}^{\mathrm{i}\varphi}$ and equate $zw$ to 1).

**1.6 Examples (complex numbers).**

- $1+\mathrm{i}$ has modulus $\sqrt{2}$ and lies in the first quadrant, on the line $y = x$. Hence $1 + \mathrm{i}$ has polar form $\sqrt{2}\mathrm{e}^{\mathrm{i}\pi/4}$.
- $(1 + \mathrm{i})^{-1} = (1/\sqrt{2})\mathrm{e}^{-\mathrm{i}\pi/4}$, from above. Alternatively, the inverse can be obtained in Cartesian form by writing

$$\frac{1}{1+\mathrm{i}} = \frac{1}{1+\mathrm{i}}\left(\frac{1-\mathrm{i}}{1-\mathrm{i}}\right) = \frac{1-\mathrm{i}}{2}.$$

- To compute quotients it is usually best to employ polars. For example,

$$\frac{-1+\sqrt{3}\mathrm{i}}{1-\mathrm{i}} = \frac{2\mathrm{e}^{\mathrm{i}2\pi/3}}{\sqrt{2}\mathrm{e}^{-\pi\mathrm{i}/4}} = \sqrt{2}\mathrm{e}^{(2\pi/3-(-\pi/4))\mathrm{i}} = \sqrt{2}\mathrm{e}^{7\pi\mathrm{i}/12}.$$

**1.7 Roots of polynomials.** The equation $x^2+1 = 0$ has no real solutions. This fact indicates the inadequacy of the real numbers as a setting for solving real polynomial equations. In $\mathbb{C}$, the equation $z^2+1 = 0$ has roots $\pm\mathrm{i}$. In Chapter 13 we shall prove the important **Fundamental theorem of algebra**, asserting that a polynomial of degree $n$ with coefficients in $\mathbb{C}$ has a full complement of $n$ roots (not neccessarily distinct, of course).

There are certain polynomials which occur repeatedly in complex analysis, and you are recommended to become very familiar with these polynomials and the location of their roots.

Let us consider the equation $z^n = 1$. Write $z = r\mathrm{e}^{\mathrm{i}\theta}$. Certainly we must have $r = 1$, because $|z|^n = |z^n|$. Also, from 1.5,

$$\begin{aligned} z^n = 1 &\iff r^n\mathrm{e}^{\mathrm{i}n\theta} = 1 \\ &\iff r = 1 \text{ and } \cos n\theta + \mathrm{i}\sin n\theta = 1. \end{aligned}$$

The results in 1.3 show that the distinct roots of the equation $z^n = 1$ are given by

$$\mathrm{e}^{2k\pi\mathrm{i}/n} \quad (k = 0, \ldots, n-1).$$

These numbers are known as the $n$th **roots of unity**. Observe that the roots of $z^n = 1$ lie at the vertices of a regular $n$-gon centred at 0 and with one vertex at 1. There is a single real root, namely $z = 1$, if $n$ is odd and precisely two real roots, namely $\pm 1$, if $n$ is even. The case $n = 6$ is illustrated in Fig. 1.3.

Two special cases are worthy of particular note. We have

$$z^4 = 1 \iff (z^2-1)(z^2+1) = 0 \iff z = \pm 1 \text{ or } z = \pm\mathrm{i}$$

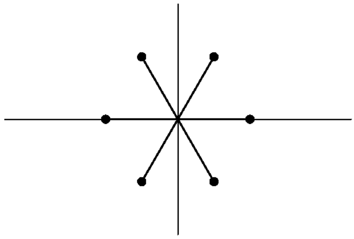


**Figure 1.3** The sixth roots of unity

and

$$z^3 = 1 \iff z = 1 \text{ or}$$
$$z = \omega := e^{2\pi i/3} = \frac{-1 + i\sqrt{3}}{2} \text{ or}$$
$$z = \omega^2 := e^{4\pi i/3} = \frac{-1 - i\sqrt{3}}{2}.$$

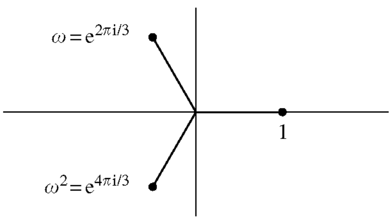



**Figure 1.4** The cube roots of unity

The **geometric identity**

$$(1 - z)(1 + z + \cdots + z^k) = (1 - z^{k+1})$$

is valid for all $z \in \mathbb{C}$ and for all natural numbers $k$. We obtain it by multiplying out the left-hand side and noticing that all but the two terms on the right-hand side cancel out. Taking $k = 2$ we obtain

$$(1 - z)(1 + z + z^2) = (1 - z^3).$$

From this we see that the roots of the equation $1 + z + z^2 = 0$ are exactly the non-real cube roots of unity, that is, $\omega$ and $\omega^2$.

As a further example, consider the equation $1+z^2+z^4=0$. We have (from the geometric identity with $z^2$ in place of $z$ and with $k=3$)

$$1+z^2+z^4=\frac{1-z^6}{1-z^2} \quad (z\neq\pm1).$$

Consequently the given equation has solutions

$$e^{\pi i/3}, \quad e^{2\pi i/3}, \quad e^{-\pi i/3}, \quad e^{-2\pi i/3}$$

—the sixth roots of unity, excluding $\pm1$.

## Conjugation, modulus, and inequalities

**1.8 Complex conjugation.** Given $z=x+iy$, the **complex conjugate** of $z$ is defined to be $\overline{z}:=x-iy$ or, in polar form, $\overline{z}:=re^{-i\theta}$. In the Argand diagram, $\overline{z}$ is the reflection of $z$ in the real axis. As examples, we note that $\overline{i}=-i$ and $\overline{\omega}=\omega^2$ (recall Fig. 1.4). The following identities hold for all $z$ and $w$ in $\mathbb{C}$:

(1) $\overline{\overline{z}}=z$;

(2) $2\operatorname{Re}z=z+\overline{z}$ and $2i\operatorname{Im}z=z-\overline{z}$;

(3) $\overline{z+w}=\overline{z}+\overline{w}$;

(4) $\overline{zw}=\overline{z}\,\overline{w}$;

(5) $|\overline{z}|=|z|$;

(6) $|z|^2=z\overline{z}$.

The formulae in (1)–(3) follow immediately from the Cartesian representations of $z$ and $w$. Formulae (4) and (5) come directly from the product formula in 1.5. The formula in (6) can be derived in various ways. Perhaps the simplest is just to note that

$$z\overline{z}=(x+iy)(x-iy)=x^2+y^2=|z|^2.$$

We have already seen that $|zw|=|z|\,|w|$. In general, $|z+w|\neq|z|+|w|$. However, important inequalities link modulus and addition.

**1.9 Inequalities.** For all $z$ and $w$ in $\mathbb{C}$,

(1) $|\operatorname{Re}z|\leqslant|z|$ and $|\operatorname{Im}z|\leqslant|z|$;

(2) $|z+w|\leqslant|z|+|w|$ (the **triangle inequality**);

(3) $|z+w|\geqslant\big||z|-|w|\big|$.

**Proof** (1) is immediate since $|z|^2 = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2$ and $|z| \geqslant 0$. To prove (2), observe that

$$\begin{aligned}
|z+w|^2 &= (z+w)(\overline{z}+\overline{w}) && \text{(by 1.8(3) \& (6))}\\
&= |z|^2 + |w|^2 + (w\overline{z} + z\overline{w}) && \text{(by 1.8(6))}\\
&= |z|^2 + |w|^2 + 2\operatorname{Re}(z\overline{w}) && \text{(by 1.8(1) \& (2))}\\
&\leqslant |z|^2 + |w|^2 + 2\,|z\overline{w}| && \text{(by (1))}\\
&= |z|^2 + |w|^2 + 2\,|z|\,|w| && \text{(by 1.8(5) \& 1.5)}\\
&= (|z| + |w|)^2.
\end{aligned}$$

Since $|z+w| \geqslant 0$ and $|z| + |w| \geqslant 0$ we deduce (2).

Now consider (3). For real numbers $a$ and $b$, the inequality $|a| \leqslant b$ holds if and only if $a \leqslant b$ and $-a \leqslant b$ (and, necessarily, $b \geqslant 0$). Hence (3) is satisfied provided the two inequalities $|z+w| \geqslant |z| - |w|$ and $|z+w| \geqslant |w| - |z|$ hold. But by (2) we have

$$\begin{aligned}
|z| &= |z+w-w| \leqslant |z+w| + |-w| = |z+w| + |w|,\\
|w| &= |z+w-z| \leqslant |z+w| + |-z| = |z+w| + |z|,
\end{aligned}$$

so (3) follows. □

All of the inequalities in (1)–(3) concern *complex* numbers but are between *real* numbers ($|z|$, $\operatorname{Re} z$, etc.). It is important to realize that no meaning has been assigned to an inequality $z \leqslant w$ between complex numbers $z$ and $w$. Indeed, it is not possible to define an ordering on $\mathbb{C}$ in which any two elements are comparable and which is compatible with the arithmetic operations in the way that the ordering on $\mathbb{R}$ is (see Exercise 1.14). Whenever inequalities appear henceforth, the quantities involved are assumed to be real. Thus $w \geqslant 0$ means that $w$ is a real number and that $w$ is also non-negative. In complex analysis, abuse of inequalities is perhaps the most common type of error perpetrated by beginners. You have been warned!

**1.10 Functions.** Formally, given $S \subseteq \mathbb{C}$, a mapping $f\colon S \to \mathbb{C}$ which assigns to each $z \in S$ a unique complex number $f(z)$ is called a **(complex-valued) function** on $S$. Being explicit about the definition of such a fundamental notion is not mere pedantry. Rather, this is an opportune point at which to emphasize the inherent 'one-valuedness' of a function. Later we deal also with what we call multifunctions: a **multifunction** is a rule assigning a non-empty subset of $\mathbb{C}$ (finite or infinite) to each element of its domain set $S$. An example, and one which underlies all other important multifunctions, comes from the argument 'function' which associates to $z$ the infinite set $\{\,\theta \in \mathbb{R} : z = |z|\,\mathrm{e}^{\mathrm{i}\theta}\,\}$.

Strictly, given a function $f$, we should distinguish between $f$ (the mapping) and $f(z)$ (the image of the point $z$ under $f$). However, where it would be cumbersome to do otherwise, we allow $f(z)$ to denote the function, and write, for example, '$z^2$' in place of 'the function $f$ defined by $f(z) = z^2$'. We also adopt the notation $z \mapsto w = f(z)$ to indicate that $z$ is mapped to $w$ by $f$.

Just as we can write a complex number in terms of its real and imaginary parts, we can express any complex-valued function $f$ in terms of real-valued functions, as follows:

$$f = u + \mathrm{i}v, \quad \text{where } u(z) = \mathrm{Re}\,\big(f(z)\big) \text{ and } v(z) = \mathrm{Im}\,\big(f(z)\big).$$

Sometimes it is profitable to link the study of functions of a complex variable $z = x + \mathrm{i}y$ to the study of functions of $(x, y)$, with $x$ and $y$ real variables. In this situation we write $f(x, y)$ in place of $f(z)$ and likewise for $u$ and $v$ (so regarding $\mathbb{C}$ as being identified with $\mathbb{R}^2$). Consider, for example, $f(z) = z^2$. Then $f(z) = x^2 - y^2 + 2\mathrm{i}xy$, so that $u(x, y) = x^2 - y^2$ and $v(x, y) = 2xy$.

## Exercises

**Exercises from the text.** Verify properties (1)–(3) in 1.8.

1.1 (a) Express each of the following in the form $r\mathrm{e}^{\mathrm{i}\theta}$:

(i) $\mathrm{i}^3$, (ii) $1 - \mathrm{i}$, (iii) $\sqrt{2}(1 + \mathrm{i})$, (iv) $\sqrt{3} - \mathrm{i}$, (v) $2 - 2\sqrt{3}\mathrm{i}$.

(b) Express each of the following in the form $x + \mathrm{i}y$ $(x, y \in \mathbb{R})$:

(i) $\mathrm{e}^{\pi\mathrm{i}/4}$, (ii) $5\mathrm{e}^{-\pi\mathrm{i}}$, (iii) $2\mathrm{e}^{3\pi\mathrm{i}/2}$, (iv) $\mathrm{e}^{4\pi\mathrm{i}/3}$, (v) $\mathrm{e}^{7\pi\mathrm{i}/6}$.

1.2 Express each of the following in polar form:

(i) $(1-\mathrm{i})(-1-\mathrm{i})$, (ii) $(1-\mathrm{i})^{-1}$, (iii) $(\sqrt{3}-\mathrm{i})/(1+\mathrm{i})$, (iv) $(1+\sqrt{3}\mathrm{i})^3$.

1.3 Express in terms of $r$, $\theta$ the following equations, where $z = r\mathrm{e}^{\mathrm{i}\theta}$:

(i) $|z^2| = 4$, (ii) $|z^2 - 1| = 1$, (iii) $\arg(2z) = 2\pi/3$,
(iv) $\arg(\mathrm{i}z) = \pi/4$, (v) $\arg z^2 = \pi/2$.

Simplify the answers as far as you can.

1.4 Evaluate, for $n = 1, 2, 3, \ldots,$

$$\text{(i) } \mathrm{i}^n, \qquad \text{(ii) } \left(\frac{1-\mathrm{i}}{1+\mathrm{i}}\right)^n, \qquad \text{(iii) } (1+\mathrm{i})^n + (1-\mathrm{i})^n.$$

(Hint: use polar representations for the powers.)

1.5 Without using the binomial expansion, show that

$$(\sqrt{3}+\mathrm{i})^n + (\sqrt{3}-\mathrm{i})^n$$

is real for any positive integer $n$.

1.6 Evaluate $\sum_{k=0}^{n} \mathrm{e}^{\mathrm{i}k\theta}$. Deduce that

$$-1 + 2\sum_{k=1}^{n} \cos k\theta = \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta} \qquad \text{for } \theta \neq 2m\pi \quad (m \in \mathbb{Z}).$$

Find a similar expression for $\sum_{k=1}^{n} \sin k\theta$.

1.7 Find the solutions of the equation $\cos n\theta + \mathrm{i}\sin n\theta = -1$ $(\theta \in \mathbb{R})$. Hence find all roots in $\mathbb{C}$ of the following equations:

$$\text{(i) } z^3 + 1 = 0, \qquad \text{(ii) } z^4 + 1 = 0, \qquad \text{(iii) } z^6 + 1 = 0.$$

Illustrate your answers geometrically.

1.8 Find all solutions in $\mathbb{C}$ of the following equations:

(i) $1 + z + \cdots + z^7 = 0$,

(ii) $(1-z)^6 = (1+z)^6$ (hint: do not multiply out!),

(iii) $1 - z + z^2 = 0$,

(iv) $1 - z^2 + z^4 - z^6 = 0$.

(Hint: recall the examples in 1.7.)

1.9 Let $\alpha$ be such that $\alpha^3 = -1$ and $\alpha \neq -1$. Evaluate $(\alpha^2(\alpha-1)^2)^{-1}$. (Hint: it is not necessary to find the possible values of $\alpha$ explicitly.)

1.10 Prove that, for $z \in \mathbb{C}$,

$$|z| \leqslant |\mathrm{Re}\, z| + |\mathrm{Im}\, z| \leqslant \sqrt{2}\,|z|\,.$$

Give examples to show that either inequality may be an equality.

1.11 Let $z, w \in \mathbb{C}$. Prove that

$$|z + \mathrm{i}w|^2 + |w + \mathrm{i}z|^2 = 2(|z|^2 + |w|^2).$$

1.12 Use the results in 1.8 to prove that, for $z$ and $w$ in $\mathbb{C}$,

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

Deduce that, if $|z|<1$ and $|w|<1$,

$$\left|\frac{z-w}{1-\overline{z}w}\right| < 1.$$

1.13 Let $z$ and $w$ be complex numbers with $z \neq w$.

(i) Prove that

$$\mathrm{Re}\left(\frac{w+z}{w-z}\right) = \frac{|w|^2-|z|^2}{|w-z|^2}.$$

(ii) Let $z = r\mathrm{e}^{\mathrm{i}\theta}$ and $w = R\mathrm{e}^{\mathrm{i}\varphi}$, with $0<r<R$. By writing $|w-z|^2$ as $(w-z)(\overline{w}-\overline{z})$ prove that

$$|w-z|^2 = R^2 - 2Rr\cos(\theta-\varphi) + r^2.$$

Deduce that

$$\mathrm{Re}\left(\frac{w+z}{w-z}\right) = \frac{R^2-r^2}{R^2-2Rr\cos(\theta-\varphi)+r^2}.$$

(These formulae for the **Poisson kernel** are needed in Chapter 23.)

1.14 The usual order relation $>$ on $\mathbb{R}$ satisfies

(a) $x \neq 0$ implies $x>0$ or $-x>0$, but not both, and

(b) $x, y > 0$ implies $x+y>0$ and $xy>0$.

Show that there does not exist a relation $>$ on $\mathbb{C}$ satisfying (a) and (b). (Hint: consider i.)

1.15 Find the real and imaginary parts of the following functions as functions of $x$ and $y$:

$$\text{(i) } z^3, \quad \text{(ii) } (z+z^{-1}) \ \ (z\neq 0), \quad \text{(iii) } 1/(1-z) \ \ (z \neq 1).$$

# 2 Geometry in the complex plane

This chapter introduces various curves and regions in the complex plane which occur repeatedly later on, and presents a variety of ways of describing these. One reason that such descriptions are important is that they assist in finding well-behaved mappings of one region of the complex plane onto another—a key technique in the application of complex variable methods to physical problems. We study mappings in Chapter 8 and also in Chapter 23, where we hint at applications. Here we investigate some particular mappings—the Möbius transformations—and their effect on lines and circles. Throughout this chapter we use geometrical ideas wherever possible rather than taking refuge in decompositions into real and imaginary parts.

## Lines and circles

**2.1 Measuring distances.** We noted in 1.3 that the modulus of $z = x + \mathrm{i}y$ has a geometric interpretation. It gives the distance of the point $(x, y)$ from the origin of coordinates, $(0, 0)$. More generally, $|z - w|$ is the distance between points $z$ and $w$ in the Argand diagram. Geometrically, the triangle inequality in 1.9(2) asserts that the length of one side of a triangle in the plane does not exceed the sum of the lengths of the other two sides.

Many sets in the plane can be geometrically described in terms of distances. For example, the points strictly above the real axis are exactly those which are closer to i than they are to $-\mathrm{i}$ (Fig. 2.1). To convert such descriptions into symbolic form we frequently make use of the fact that $|z - a| < r$, $|z - a| = r$, or $|z - a| > r$ according as $z$ is at a distance less than, equal to, or greater than, $r$ from $a$.

**2.2 Equations for line segments and lines.** For $\alpha$ and $\beta$ in $\mathbb{C}$, the **line segment** with endpoints $\alpha$ and $\beta$ is given by

$$[\alpha, \beta] := \{\, (1 - t)\alpha + t\beta \,:\, 0 \leqslant t \leqslant 1 \,\}.$$

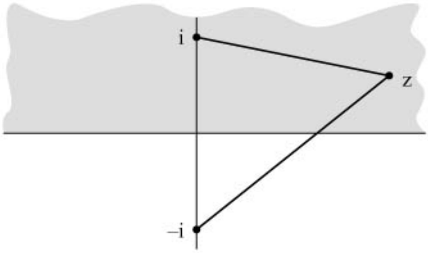



**Figure 2.1** Points closer to i than to $-\mathrm{i}$

If we consider $\{ (1-t)\alpha + t\beta \ : \ t \in \mathbb{R}\}$ we obtain all the points on the unique line through $\alpha$ and $\beta$, assuming these points are distinct.

To indicate the variety of different descriptions that may be available for the same linear locus, we note that the **real axis** is given by any of the equations:

- $\operatorname{Im} z = 0$;
- $z = \overline{z}$;
- $|z - \mathrm{i}| = |z + \mathrm{i}|$ or, more generally, $|z - \alpha| = |z - \overline{\alpha}|$ (where $\operatorname{Im} \alpha \neq 0$).

The first of these may be taken as the definition. The second is clearly equivalent to it, by 1.8(2). The conjugate $\overline{\alpha}$ of $\alpha$ is simply the reflection of $\alpha$ in the real axis; the last equation characterizes the real axis as the set of points equidistant from $\alpha$ and $\overline{\alpha}$.

The **imaginary axis** can be similarly described. In particular it is given by either the equation $\operatorname{Re} z = 0$ or the equation $|z - 1| = |z + 1|$.

The **perpendicular bisector** of the line segment joining distinct points $\alpha$ and $\beta$ in $\mathbb{C}$ is given by $|z - \alpha| = |z - \beta|$. The equation of any line can, for suitable $\alpha$ and $\beta$, be expressed in this convenient form.

**2.3 Equations for circles.** The **circle** centre $a \in \mathbb{C}$ and radius $r > 0$ is the locus of points at distance $r$ from $a$ and so has equation $|z - a| = r$. There is another form of equation which also specifies a circle. We reveal the full benefits of this representation in Chapter 8. Let $\alpha, \beta \in \mathbb{C}$ with $\alpha \neq \beta$ and let $\lambda \in \mathbb{R}$, with $\lambda > 0$ and $\lambda \neq 1$. We claim that the equation

$$\left| \frac{z - \alpha}{z - \beta} \right| = \lambda$$

represents a circle. Denote $\alpha$ by $A$ and $\beta$ by $B$, and the variable point $z$ by $P$. Then the given equation specifies the locus of points $P$ for which the ratio $AP : PB$ has the constant value $\lambda$. This locus can be shown geometrically to be a circle, known as a **circle of Apollonius**; see Fig. 2.2. However, it is simpler to switch to Cartesian coordinates (a strategy we stress is usually best avoided in complex analysis). Put $\alpha = \alpha_1 + i\alpha_2$, $\beta = \beta_1 + i\beta_2$, $z = x + iy$. The given equation can be rewritten as

$$|z-\alpha|^2 = \lambda^2 |z-\beta|^2,$$

which can be expressed in the form

$$(x-\alpha_1)^2 + (y-\alpha_2)^2 = \lambda^2(x-\beta_1)^2 + \lambda^2(y-\beta_2)^2.$$

This simplifies to

$$\left(x - \frac{\alpha_1 - \lambda^2\beta_1}{1-\lambda^2}\right)^2 + \left(y - \frac{\alpha_2 - \lambda^2\beta_2}{1-\lambda^2}\right)^2 = K,$$

where $K$ is a positive constant. Thus we do indeed have the equation of a circle. Conversely, every circle can be represented in this way. This can be proved either analytically (see 2.10) or geometrically.

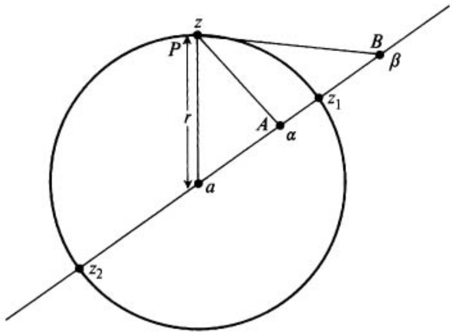



**Figure 2.2** The circle of Apollonius

**2.4 Equations for circular arcs.** We have already considered line segments with endpoints $\alpha$ and $\beta$. We now want to consider circular arcs joining $\alpha$ and $\beta$. We need to use angles to describe such arcs. As in Chapter 1, we use $\arg z$ to denote any choice of angle $\theta$ such that $z = |z|\,e^{i\theta}$; we allow $z = 0$ here and assign $\theta$ arbitrarily in this case. The following facts come from the polar representations of products and quotients:

$$\arg(z_1 z_2) = \arg z_1 + \arg z_2 \pmod{2\pi},$$
$$\arg(z_1/z_2) = \arg z_1 - \arg z_2 \pmod{2\pi}.$$

Let $P$ be a variable point on a circular arc with endpoints $A$ and $B$ in $\mathbb{C}$. Then, by elementary geometry, the angle $\angle APB$ is a constant, $\mu$ say. Write $P$ as $z$, $A$ as $\alpha$, and $B$ as $\beta$. Let $\arg(z-\alpha) = \theta$ and $\arg(z-\beta) = \phi$, as in Fig. 2.3. We have $\mu = \theta - \phi$. Hence the arc $APB$ has equation

$$\arg(z-\alpha) - \arg(z-\beta) = \mu \pmod{2\pi},$$

that is,

$$\arg\left(\frac{z-\alpha}{z-\beta}\right) = \mu \pmod{2\pi}.$$

(Strictly, we should exclude $z = \alpha, \beta$ here, but it will be convenient to regard these points as included in the locus.) The case $\mu = \pi$ gives the degenerate case when the arc through $\alpha$ and $\beta$ is the line segment joining these points.

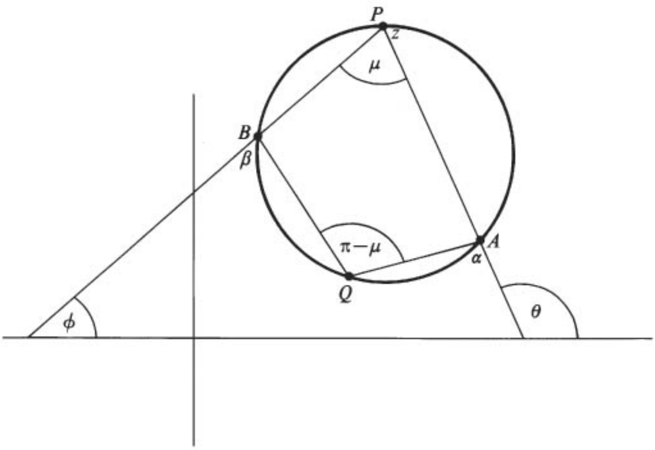



**Figure 2.3** Circular arcs

Similarly, the equation of the arc $AQB$ in Fig. 2.3 is (note the signs!)

$$\arg\left(\frac{z-\alpha}{z-\beta}\right) = -(\pi-\mu) \pmod{2\pi}.$$

Although it is worthwhile to have these representations, they need to be used carefully, since it is easy to make sign errors.

**2.5 Example (circular arcs).** Consider $S := \{ z \in \mathbb{C} : |z| \leqslant 1, \operatorname{Im} z \leqslant 0 \}$. This is the semicircular area shown in Fig. 2.4. It is made up of the arcs

$$\arg\left(\frac{z+1}{z-1}\right) = \mu \pmod{2\pi} \quad \text{for } \pi/2 \leqslant \mu \leqslant \pi.$$

The boundary of $S$ is given by the line segment $[-1, 1]$ (the equation above with $\mu = \pi$) and the arc through $-1, -\mathrm{i}$, and $1$ (take $\mu = \pi/2$).

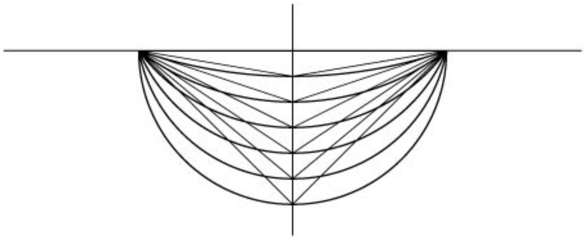


**Figure 2.4** Example 2.5

**2.6 An assortment of subsets of $\mathbb{C}$.** Many subsets of $\mathbb{C}$ which appear repeatedly later on have boundaries which are curves of the types described above. The usage 'open', 'closed', and 'region' is compatible with the technical meanings these terms acquire in Chapter 3.

- **Discs** The **open disc** centre $a \in \mathbb{C}$ and radius $r > 0$ is defined to be

$$\mathrm{D}(a;r) := \{ z \in \mathbb{C} : |z-a| < r \}.$$

The **closed disc** centre $a$ and radius $r > 0$ is

$$\overline{\mathrm{D}}(a;r) := \{ z \in \mathbb{C} : |z-a| \leqslant r \}.$$

It is the union of $\mathrm{D}(a;r)$ with its bounding circle $|z-a| = r$. The **punctured disc** centre $a \in \mathbb{C}$ and radius $r > 0$ is

$$\mathrm{D}'(a;r) := \{ z \in \mathbb{C} : 0 < |z-a| < r \}.$$

Since $z \neq a$ if and only if $0 < |z-a|$, we have $\mathrm{D}'(a;r) = \mathrm{D}(a;r) \setminus \{a\}$.

- **Annuli** Any set of the form

$$\{ z \in \mathbb{C} : s < |z-a| < r \} \quad (0 \leqslant s < r) \quad \text{or}$$
$$\{ z \in \mathbb{C} : s < |z-a| \} \quad (0 \leqslant s)$$

is called an (open) **annulus**. The case $s = 0$ in the first of these gives the punctured disc $\mathrm{D}'(a; r)$.

- **Half-planes** The **open upper half-plane** is $\Pi^+ := \{ z \in \mathbb{C} : \operatorname{Im} z > 0 \}$. The **closed upper half-plane** has the bounding line $\operatorname{Im} z = 0$ included. To obtain this, simply change $>$ to $\geqslant$ in the definition of $\Pi^+$. The **open lower half-plane** is $\Pi^- := \{ z \in \mathbb{C} : \operatorname{Im} z < 0 \}$. Other half-planes are defined analogously.
- **Sectors** $S_{\alpha,\beta} := \{ z \in \mathbb{C} : 0 \neq z = |z| \mathrm{e}^{\mathrm{i}\theta} \in \mathbb{C} \text{ with } \alpha < \theta < \beta \}$ $(\alpha < \beta)$ is a **sector**. We have a **quadrant** if $\beta - \alpha = \pi/2$ and a half-plane if $\beta - \alpha = \pi$. For example, $\Pi^+$ is $\{ z \in \mathbb{C} : 0 \neq z = |z| \mathrm{e}^{\mathrm{i}\theta} \in \mathbb{C} \text{ with } 0 < \theta < \pi \}$.

**Figure 2.5** An annulus, a half-plane, and a sector

## The extended complex plane and the Riemann sphere

We next introduce an ingenious device, due to Riemann, which will allow us to treat lines and circles, and half-lines and circular arcs, in a unified way.

**2.7 The Riemann sphere and the extended complex plane.** Let us regard $\mathbb{C}$ as embedded in Euclidean space $\mathbb{R}^3$ by identifying $z = x + \mathrm{i}y$ with $(x, y, 0)$. Now let

$$\Sigma := \{ (x, y, u) \in \mathbb{R}^3 : x^2 + y^2 + \left(u - \tfrac{1}{2}\right)^2 = \tfrac{1}{4} \}.$$

This is a sphere (the **Riemann sphere**). It touches the plane $\mathbb{C}$ at $S := (0,0,0)$. Stereographic projection allows us to set up a one-to-one correspondence between the points of $\mathbb{C}$ and the points of $\Sigma$, excluding $N$, the north pole of $\Sigma$. Geometrically, the line from any point $z$ of $\mathbb{C}$ to $N$ cuts $\Sigma \setminus \{N\}$ in precisely one point $z'$, and, for every point $z'$ of $\Sigma \setminus \{N\}$, the line through $N$ and $z'$ meets the plane $\mathbb{C}$ in a unique point $z$. The irritation of the north pole being 'left out in the cold' can be removed: we add to $\mathbb{C}$ an extra point $\infty \notin \mathbb{C}$ and define the **extended complex plane** $\widetilde{\mathbb{C}}$ to be $\mathbb{C} \cup \{\infty\}$. We then have a natural correspondence between $\widetilde{\mathbb{C}}$ and $\Sigma$, which is given analytically by

$$\mathbb{C} \ni z = x + \mathrm{i}y = r\mathrm{e}^{\mathrm{i}\theta} \longleftrightarrow z' = (x(1+r^2)^{-1}, y(1+r^2)^{-1}, r^2(1+r^2)^{-1}),$$
$$\infty \longleftrightarrow (0,0,1).$$

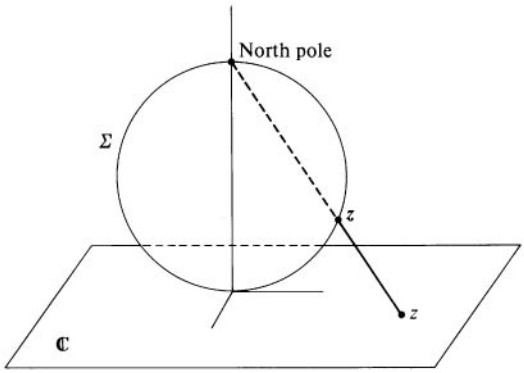



**Figure 2.6** Stereographic projection

It might at first sight seem curious to have a *single* 'point at infinity' adjoined to $\mathbb{C}$—after all, separate symbols $\infty$ and $-\infty$ are associated with the real line. However, it should be clear that, if we let $z = r\mathrm{e}^{\mathrm{i}\alpha}$ (with $\alpha$ fixed) and allow $r$ to become arbitrarily large, then $z'$ will approach $N$ regardless of the value of $\alpha$: informally stated, 'all roads lead to $\infty$'!

**2.8 Arithmetic in the extended plane.** We often wish to make use of the algebraic structure of $\mathbb{C}$ and it is convenient to extend this, with some provisos, to $\widetilde{\mathbb{C}}$. We adopt the following conventions when working with $\widetilde{\mathbb{C}}$:

$$\begin{aligned} &a \pm \infty = \pm\infty + a = \infty, && a/\infty = 0 \quad \text{for all } a \in \mathbb{C}, \\ &a \cdot \infty = \infty \cdot a = \infty, && a/0 = \infty \quad \text{for all } a \in \mathbb{C} \setminus \{0\}, \\ &\infty + \infty = \infty \cdot \infty = \overline{\infty} = \infty. && \end{aligned}$$

We are allowed to divide a complex number $a$ by zero!—so long as $a \neq 0$. Note that $\infty - \infty$ and $\infty/\infty$ are not assigned meanings.

**2.9 Circlines.** The merit of moving to $\widetilde{\mathbb{C}}$, or equivalently to $\Sigma$, becomes clearer when we consider lines and circles. For this purpose we shall work with lines in $\widetilde{\mathbb{C}}$ rather than $\mathbb{C}$, by regarding $\infty$ as being adjoined to any line in $\mathbb{C}$. Under the stereographic correspondence above, circles on $\Sigma$ which pass through the north pole $N$ project down onto lines in $\widetilde{\mathbb{C}}$.

Now take any circle drawn on $\Sigma$ parallel to the horizontal plane $u = 0$, necessarily with centre on the vertical axis $x = y = 0$. Certainly, we obtain a circle in $\mathbb{C}$ (with centre at 0) when this is projected down to $\mathbb{C}$. The smaller the radius of the original circle, the larger the radius of its projection. Extending this, it can be shown that any circle on $\Sigma$ which does not pass through $N$ projects onto a circle in $\mathbb{C}$, and that every circle in $\mathbb{C}$ arises in this manner.

With this perspective, it is now natural to regard lines (in $\widetilde{\mathbb{C}}$) as 'circles through infinity', and to adopt the collective name **circline** for a circle or straight line in $\widetilde{\mathbb{C}}$. Pulling together our conclusions from 2.2 and 2.3, we see that the equation

$$\left|\frac{z-\alpha}{z-\beta}\right| = \lambda \quad (\lambda > 0)$$

represents a circline, and that every circline can be represented in this form. The circline is a line if $\lambda = 1$ and a circle (in $\mathbb{C}$) otherwise.

By a **segment of a circline** (or simply an **arc**) we mean the arc of a circle joining points $\alpha$ and $\beta$ in $\mathbb{C}$, the line segment joining points $\alpha$ and $\beta$ in $\mathbb{C}$, or a half-line from $\alpha \in \mathbb{C}$ to $\infty$ (in any direction). Such a half-line (also called a **ray**) has an equation of the form $\arg(z - \alpha) = \mu$.

**2.10 Inverse points.** Consider again the circline given by the equation

$$\left|\frac{z-\alpha}{z-\beta}\right| = \lambda \quad (\lambda > 0).$$

We now investigate the geometric significance of the points $\alpha$ and $\beta$, which are known as **inverse points** with respect to the circline.

First consider the case $\lambda = 1$, so that we have a line, $L$ say. Then the points $\alpha$ and $\beta$ are reflections of each other in $L$.

Now let $\lambda \neq 1$. There are exactly two points on the circle which are collinear with $\alpha$ and $\beta$: these are the points $z_1$ and $z_2$ which satisfy

$$z_1 - \alpha = \lambda(z_1 - \beta) \quad \text{and} \quad z_2 - \alpha = -\lambda(z_2 - \beta).$$

These lie at opposite ends of a diameter. Writing the equation of the circle in the form $|z - a| = r$ we have

$$a = \frac{1}{2}(z_1 + z_2) \quad \text{and} \quad r = \frac{1}{2}\left|z_1 - z_2\right|.$$

Hence

$$\alpha - a = \frac{1}{2}\big((\alpha - z_1) + (\alpha - z_2)\big) = \frac{1}{2}\lambda\big((\beta - z_1) - (\beta - z_2)\big) = \frac{1}{2}\lambda(z_1 - z_2)$$

and, likewise,

$$\lambda(\beta - a) = \frac{1}{2}(z_1 - z_2).$$

Consequently (remember 1.8(6))

$$(\alpha - a)\overline{(\beta - a)} = \frac{1}{4}(z_1 - z_2)\overline{(z_1 - z_2)} = r^2.$$

So points $\alpha$ and $\beta$ in $\mathbb{C}$ are inverse points with respect to the circle $|z - a| = r$ if and only if they satisfy $(\alpha - a)\overline{(\beta - a)} = r^2$. Note that we must always have one of $\alpha$ and $\beta$ inside the circle and the other outside. We shall also regard $\alpha = a$ and $\beta = \infty$ as a pair of inverse points for $|z - a| = r$.

## 2.11 Examples (inverse-point representation of circles).

- Consider the circle given by $|z - 3| = 2$. This has the segment $[1, 5]$ of the real axis as a diameter. The real numbers $\alpha$ and $\beta$ are inverse points if and only if $(\alpha - 3)(\beta - 3) = 4$. A possible choice is $\alpha = 7$ and $\beta = 4$. So the equation of the circle can be written in the form $|z - 7| / |z - 4| = \lambda$. To find $\lambda$ we note that $z = 5$ lies on the circle. This gives $\lambda = 2/1 = 2$.
- Consider the **unit circle**, $|z| = 1$. A typical pair of inverse points for this circle is $\alpha$ and $1/\overline{\alpha}$, giving as the inverse-point representation of the circle

$$\left|\frac{z - \alpha}{z - 1/\overline{\alpha}}\right| = \left|\frac{1 - \alpha}{1 - 1/\overline{\alpha}}\right|,$$

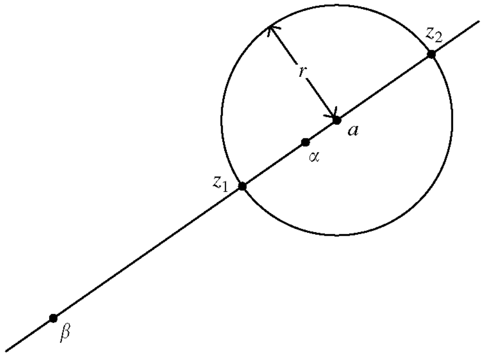



**Figure 2.7** Circles from inverse-point representation

that is,

$$\left|\frac{z-\alpha}{\overline{\alpha}z-1}\right| = 1.$$

Here we have used the fact that 1 lies on the circle and that $|\alpha - 1| = |\overline{\alpha} - 1|$.

- Consider the circle with equation $|z+1| = \lambda|z|$, where $\lambda \neq 1$. It has $-1$ and 0 as inverse points, and hence has a diameter lying on the real axis. The ends of this diameter are at the solutions of $z + 1 = \pm\lambda z$, that is, at $(\pm\lambda - 1)^{-1}$. The centre is at $(\lambda^2 - 1)^{-1}$ and the radius is $\lambda\left|\lambda^2 - 1\right|^{-1}$.

**2.12 Coaxal circles.** For distinct fixed points $\alpha$ and $\beta$ in $\mathbb{C}$, we have, as $\lambda$ and $\mu$ vary, two families of circles:

- $C_1(\alpha, \beta)$: circles

$$\left|\frac{z-\alpha}{z-\beta}\right| = \lambda,$$

having $\alpha$ and $\beta$ as inverse points. Here $\lambda = 1$ is allowed, so that the straight line bisecting the segment $[\alpha, \beta]$ is included.

- $C_2(\alpha, \beta)$: circles

$$\arg\left(\frac{z-\alpha}{z-\beta}\right) = \begin{cases} \mu \\ -(\pi - \mu) \end{cases} \pmod{2\pi},$$

through $\alpha$ and $\beta$.

Traditionally, each of the families $C_1(\alpha, \beta)$ and $C_2(\alpha, \beta)$ is said to form a **coaxal system.** Coaxal systems of circles have interesting geometric properties. For example, any member of $C_1(\alpha, \beta)$ cuts any member of $C_2(\alpha, \beta)$ orthogonally.

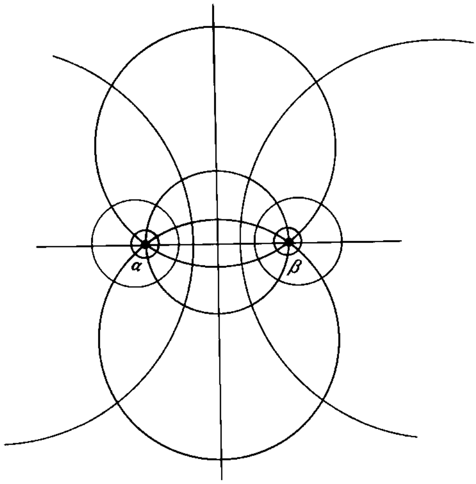



**Figure 2.8** Coaxal circles

## Möbius transformations

This section introduces a family of mappings of the extended plane onto itself which map circlines to circlines. More extensive discussions of these mappings and of their wider role in algebra and geometry can be found in [**13**] and [**18**].

**2.13 Some particular mappings.** Consider the mappings

$z \mapsto ze^{i\varphi} \quad (\varphi \in \mathbb{R})$ (anticlockwise rotation through $\varphi$),
$z \mapsto Rz \quad (R > 0)$ (stretching by a factor of $R$),
$z \mapsto z + a \quad (a \in \mathbb{R})$ (translation by $a$),
$z \mapsto 1/z$ (inversion).

It is easy to see geometrically that mappings of the first three types take straight lines to straight lines and circles to circles. For example, the real axis is mapped onto

the imaginary axis by $z \mapsto w := ze^{i\pi/2}$,
the line $\operatorname{Im} w = 1$ by $z \mapsto w := z + i$,

while the unit circle is mapped to

$$\begin{array}{lll} \text{the circle } |w| = 2 & \text{by} & z \mapsto w = 2z, \\ \text{the circle } |w-1| = 1 & \text{by} & z \mapsto w := z+1. \end{array}$$

Now consider the inversion $z \mapsto w := 1/z$, regarded as a map from $\widetilde{\mathbb{C}}$ to $\widetilde{\mathbb{C}}$. Consider the image of the line $\operatorname{Re} z = 1$, which we may describe in inverse-point form as $|z| = |z-2|$ (points $z$ equidistant from 0 and 2). We have $z = 1/w$, so the image is given by $|1/w| = |1/w - 2|$, that is, $|2w-1| = 1$. Since inversion is self-inverse, we see also that the circle $|2z-1| = 1$ maps to the line $\operatorname{Re} w = 1$ under $z \mapsto w = 1/z$. We conclude that under inversion a line may map to a circle and a circle may map to a line.

**2.14 Möbius transformations.** We now bring together the four special types of mapping—rotation, stretching, translation, and inversion—considered in 2.13. A **Möbius transformation** is a mapping of the form

$$z \mapsto w = f(z) := \frac{az+b}{cz+d} \quad (a, b, c, d \in \mathbb{C},\ ad - bc \neq 0).$$

(The excluded case $ad - bc = 0$ produces a constant mapping or one which is undefined.) We view the Möbius transformation $f$ as a mapping from $\widetilde{\mathbb{C}}$ to $\widetilde{\mathbb{C}}$: $f(-d/c) = \infty$ and $f(\infty) = a/c$, according to the arithmetic rules in 2.8. Then $f\colon \widetilde{\mathbb{C}} \to \widetilde{\mathbb{C}}$ is one-to-one and onto, with a well-defined inverse given by

$$f^{-1}\colon w \mapsto \frac{dw-b}{a-cw}.$$

Certainly the special mappings described in 2.13 are Möbius transformations. Furthermore, a general Möbius transformation may be built up from mappings of these types by composing maps in the usual way. It is easy to check that the Möbius transformations form a group under the operation of composition of maps.

**2.15 Theorem (circlines under Möbius transformations).** Let $C$ be a circline with inverse points $\alpha$ and $\beta$ (in $\mathbb{C}$) and let $f$ be a Möbius transformation. Then $f$ maps $C$ to a circline, with inverse points $f(\alpha)$ and $f(\beta)$.

**Proof** We write the equation of $C$ in inverse-point form as $|z-\alpha| \,/\, |z-\beta| = \lambda$. Suppose $w = f(z) = (az+b)/(cz+d)$, so that $z = (dw-b)/(a-cw)$. We substitute for $z$ in the equation for $C$ to obtain the image $f(C)$. It is given by

$$\left|\frac{(\alpha c + d)w - (\alpha a + b)}{(\beta c + d)w - (\beta a + b)}\right| = \lambda.$$

We may rewrite this as

$$\text{(i)} \qquad \left|\frac{w - f(\alpha)}{w - f(\beta)}\right| = \lambda \left|\frac{\beta c + d}{\alpha c + d}\right| \qquad \text{if } \alpha c + d \neq 0 \text{ and } \beta c + d \neq 0, \text{ or}$$

$$\text{(ii)} \qquad |w - f(\alpha)| = \lambda \left|\frac{\beta a + b}{\alpha c + d}\right| \qquad \text{if } \alpha c + d \neq 0 \text{ and } \beta c + d = 0, \text{ or}$$

$$\text{(iii)} \qquad |w - f(\beta)| = \lambda \left|\frac{\alpha a + b}{\beta c + d}\right| \qquad \text{if } \alpha c + d = 0 \text{ and } \beta c + d \neq 0.$$

(Note that $\alpha c + d$ and $\beta c + d$ cannot both be zero, because $ad - bc \neq 0$ by assumption.) In each case $f(C)$ is indeed a circline with $f(\alpha)$ and $f(\beta)$ as inverse points; in cases (ii) and (iii), the images of the original inverse points are the centre of the circle and the point $\infty$.

**2.16 Finding images: substitution method.** With the geometric insight provided by Theorem 2.15, we can now describe the effect of Möbius transformations on circlines without having to resort to real and imaginary parts. First we draw attention explicitly to a tactic employed in the proof of Theorem 2.15 for finding an image (in this case a circline) under a mapping (in this case a Möbius transformation). Suppose that $S \subseteq \widetilde{\mathbb{C}}$ and that we wish to find the image of $S$ under a map $f\colon z \mapsto f(z) = w \in \widetilde{\mathbb{C}}$. We seek to express $z$ in terms of $w$ (this obviously requires $f$ to be injective) and then substitute this expression for $z$ into the defining relation for $S$ to get a corresponding defining relation for $f(S)$. This strategy is used in 2.13 and in 2.17.

**2.17 Example (circlines under a Möbius transformation).** Consider $f\colon z \mapsto f(z) = w := (z-1)^{-1}$. Under this Möbius transformation the image of any circline is a circline, by 2.15. We find the images of various circlines.

- The imaginary axis is given by $|z - 1| = |z + 1|$. Substituting for $z$ in terms of $w$ we see that the image has equation

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

  which reduces to $|2w + 1| = 1$. Hence the image is the circle centre $-1/2$ and radius $1/2$.

- The real axis has i and $-$i as inverse points and hence its image is a circline with $f(\mathrm{i}) = \frac{1}{2}(-1-\mathrm{i})$ and $f(-\mathrm{i}) = \frac{1}{2}(-1+\mathrm{i})$ as inverse points. These points are conjugates of each other. Also, the image passes through $\infty = f(1)$ and so is a straight line. These facts identify the image as the real axis.
  Alternatively, but more laboriously, write the equation of the real axis as
  $$|z-\mathrm{i}| = |z+\mathrm{i}|$$
  and substitute $(w+1)/w$ for $z$ to get $|w(1-\mathrm{i})+1| = |w(1+\mathrm{i})+1|$. Equivalently (remember that $|1-\mathrm{i}| = \left|\overline{1-\mathrm{i}}\right| = |1+\mathrm{i}|$),
  $$\left|w - \tfrac{1}{2}(\mathrm{i}+1)\right| = \left|w - \tfrac{1}{2}(-\mathrm{i}+1)\right|.$$
  This represents the real axis.
  For a third method, we may argue as follows. It is easy to see graphically that the map $x \mapsto (x-1)^{-1}$ is a map of $\mathbb{R} \setminus \{1\}$ onto $\mathbb{R} \setminus \{0\}$. Also, in the extended plane, $z = \infty$ corresponds to $w = 1$ and $z = 0$ to $w = \infty$. Hence the real axis (in $\widetilde{\mathbb{C}}$) is mapped to itself.
- The circle centre 0 and radius $r$ has image given by $|w+1| = r\,|w|$. If $r = 1$ then it is the line $\operatorname{Re} z = -1/2$. If $r \neq 1$ then it is a circle with $-1$ and 0 as inverse points. For explicit identification of this circle, use the formulae in 2.11.

**2.18 Triplet representation of Möbius transformations.** If we are interested in finding the image of some circline under a Möbius transformation, one way to proceed is to exploit the fact that there is one and only one circline through any triplet of three distinct points in $\widetilde{\mathbb{C}}$. Thus in the second example in 2.17 we argued that the real axis is the unique circline through 0, 1, and $\infty$, so that its image under $z \mapsto (z-1)^{-1}$ is the unique circline through $-1$, $\infty$, and 0, *viz.* the real axis. We do not actively encourage use of triplets since later on we want to map regions rather than curves. For these the substitution method will be more effective. However triplets are valuable for *constructing* Möbius transformations. Suppose that each of $\{z_1, z_2, z_3\}$ and $\{w_1, w_2, w_3\}$ is an ordered triplet of distinct points in $\widetilde{\mathbb{C}}$. Then we claim that there is a unique Möbius transformation $f$ such that $f(z_k) = w_k$ $(k = 1, 2, 3)$ and that this is given by
$$\left(\frac{w-w_1}{w-w_3}\right)\left(\frac{w_2-w_3}{w_2-w_1}\right) = \left(\frac{z-z_1}{z-z_3}\right)\left(\frac{z_2-z_3}{z_2-z_1}\right).$$
To prove the claim, note that the map
$$g\colon z \mapsto \left(\frac{z-z_1}{z-z_3}\right)\left(\frac{z_2-z_3}{z_2-z_1}\right)$$

takes $z_1$, $z_2$, $z_3$ to 0, 1, $\infty$, respectively. Construct $h$ in the same way as $g$, to map $w_1$, $w_2$, $w_3$ to 0, 1, $\infty$. Then the composite map $f := h^{-1} \circ g$ is a Möbius transformation taking $z_k$ to $w_k$, for $k = 1, 2, 3$, and it is given by the formula above.

Finally, to prove uniqueness, it is enough to show that the only Möbius transformation $f\colon z \mapsto (az + b)/(cz + d)$ $(ad - bc \neq 0)$ taking 0, 1, $\infty$ to 0, 1, $\infty$ is the identity map $z \mapsto z$. The conditions $f(0) = 0$, $f(\infty) = \infty$, $f(1) = 1$ force in turn $b = 0$, $c = 0$, $a = d$, so that $f(z) = z$ for all $z$, as required.

## Exercises

**Exercises from the text.** [Only for those proficient in geometry: verify the claim in 2.9 that under stereographic projection any circle on $\Sigma$ maps to a circline in $\widetilde{\mathbb{C}}$.]

Verify the unproved assertions in 2.14.

2.1 What do the following equations represent geometrically? Give sketches.

(i) $|z + 2| = 6$, (ii) $|z - 3\mathrm{i}| = |z + \mathrm{i}|$,
(iii) $|\mathrm{i}z - 1| = |\mathrm{i}z + 1|$, (iv) $|z - \omega| = |z - 1|$ $(\omega = \mathrm{e}^{2\pi\mathrm{i}/3})$.

2.2 Describe geometrically the subsets of $\mathbb{C}$ specified by

(i) $\mathrm{Im}(z + \mathrm{i}) < 2$, (ii) $|z - \mathrm{i}| < |z - 1|$,
(iii) $|z + 2\mathrm{i}| \geqslant 2$, (iv) $|z - 1 + \mathrm{i}| \geqslant |z - 1 - \mathrm{i}|$,
(v) $\mathrm{Im}[(z + \mathrm{i})/(2\mathrm{i})] < 0$, (vi) $1 < \mathrm{Re}\, z \leqslant 2$,
(vii) $\mathrm{Re}\, z \neq 0$, (viii) $|z - 1| < 1$ and $|z| = |z - 2|$.

2.3 Describe geometrically the subsets of $\mathbb{C}$ specified by

(i) $|z - 1 - \mathrm{i}| > 1$, (ii) $|z + \mathrm{i}| \neq |z - \mathrm{i}|$,
(iii) $z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$ $(-\pi < \theta < \frac{\pi}{2})$, (iv) $|z - 2| > 3$ and $|z| < 2$,
(v) $\mathrm{Re}\, z < 1$ or $\mathrm{Im}(z - 1) \neq 0$, (vi) $1 < |z - 1| < 2$,
(vii) $1 < \mathrm{Im}\, z < 2$ and $\mathrm{Re} > 1$, (viii) $|z|^2 > z + \overline{z}$.

2.4 Sketch the following circlines, finding the centre and radius of those which are circles.

(i) $|z + \mathrm{i}| = |z - 3\mathrm{i}|$, (ii) $|z + 1| = 4\,|z - 1|$,
(iii) $|z - \mathrm{i}| = 2\,|z|$, (iv) $2\,|z - \mathrm{i}| = |z|$.

2.5 Find a pair of inverse points in $\mathbb{C}$ for each of the following circles and hence find an equation for each in inverse-point form.

$$\text{(i) } |z-1| = 2, \qquad \text{(ii) } |z-\mathrm{i}| = \sqrt{2}, \qquad \text{(iii) } |z-1-\mathrm{i}| = 2.$$

2.6 Let $f$ be a Möbius transformation. Let $S$ be a circline, so that $f(S)$ is a circline, by Theorem 2.15. Let $\alpha$ and $\beta$ be distinct points on $S$ and consider an arc $A$ on $S$ which has endpoints $\alpha$ and $\beta$. Describe the possible geometric forms the image $f(A)$ may take. (Hint: there is a unique point $p$ in $\widetilde{\mathbb{C}}$ such that $f(p) = \infty$; consider cases (a) $p \notin S$, (b) $p = \alpha$ or $\beta$, (c) $p \in S \setminus A$, (d) $p \in (A \setminus \{\alpha, \beta\})$.)

2.7 For each of the following sets of points find an arg equation for the arc which has the first and third points as its endpoints and which passes through the second point:

$$\text{(i) } -1, \mathrm{i}, 1, \quad \text{(ii) } -1, -\mathrm{i}, \mathrm{i}, \quad \text{(iii) } -1-\mathrm{i}, 0, 1+\mathrm{i}, \quad \text{(iv) } 0, 1+\mathrm{i}, \infty.$$

2.8 Which arcs are given by the following equations? Give sketches.

$$\text{(i) } \arg\left(\frac{z+1}{z-1}\right) = \frac{\pi}{2}, \quad \text{(ii) } \arg\left(\frac{z-1}{z+1}\right) = \frac{\pi}{2}, \quad \text{(iii) } \arg\left(\frac{z+1}{z-\mathrm{i}}\right) = \pi.$$

2.9 (a) Sketch the circles $|z-1| = \sqrt{2}$ and $|z+1| = \sqrt{2}$ and find the points $\alpha$ and $\beta$ where they intersect.

(b) Find equations for the four circular arcs with endpoints $\alpha$ and $\beta$ which are arcs of the circles in (a).

(c) Describe in terms of arg the set

$$G := \{\, z \in \mathbb{C} : |z-1| < \sqrt{2},\ |z+1| < \sqrt{2} \,\}.$$

(d) Describe in terms of arg the three sets obtained from $G$ by changing one or both of the occurrences of $<$ to $>$.

2.10 Find the image (in $\widetilde{\mathbb{C}}$) of (a) the ray $\arg z = \pi/6$, (b) the disc $\mathrm{D}(0;2)$, and (c) the line $\operatorname{Im} z = 1$ under

$$\text{(i) } z \mapsto -z, \qquad \text{(ii) } z \mapsto (1+\mathrm{i})z, \qquad \text{(iii) } z \mapsto 1/z.$$

2.11 Find the Möbius transformation mapping $0,\ 1,\ \infty$ to $1,\ 1+\mathrm{i},\ \mathrm{i}$, respectively. Under this mapping what is the image of

(i) a circular arc through $-1$ and $-\mathrm{i}$,

(ii) the line given by $\operatorname{Im} z = \operatorname{Re} z$,

(iii) the real axis,
(iv) the imaginary axis?

2.12 (a) Consider the Möbius transformation $z \mapsto w = (z-1)^{-1}$ regarded as a map from $\widetilde{\mathbb{C}}$ to $\widetilde{\mathbb{C}}$. Find the images of the half-lines $[1,\infty)$ and $(-\infty,1]$ on the real axis.
(b) Consider the Möbius transformation $z \mapsto w = (z+1)/(z-1)$ regarded as a map from $\widetilde{\mathbb{C}}$ to $\widetilde{\mathbb{C}}$. Find the images of the real intervals $[0,1]$, $[1,\infty)$ and $(-\infty,1]$. Find also the image of each of the circles

$$\text{(i) } |z| = 1, \quad \text{(ii) } \left|z+\tfrac{1}{2}\right| = 1, \quad \text{(iii) } \left|z-\tfrac{1}{2}\right| = 1.$$

2.13 Let $\alpha$ be such that $|\alpha| < 1$. Let $\phi_\alpha(z) = (z-\alpha)/(\overline{\alpha}z - 1)$. Show that $\phi_\alpha$ maps $\mathrm{D}(0;1)$ one-to-one onto $\mathrm{D}(0;1)$ and that the inverse map $\phi_\alpha^{-1}$ is $\phi_\alpha$. Prove in addition that every Möbius transformation mapping $\mathrm{D}(0;1)$ onto $\mathrm{D}(0;1)$ is of the form $\mathrm{e}^{\mathrm{i}\lambda}\phi_\alpha$, for some real constant $\lambda$ and some $\alpha \in \mathrm{D}(0;1)$.

2.14 Let $f\colon z \mapsto w = (az+b)/(cz+d)$ $(ad-bc \neq 0)$ be a Möbius transformation, other than the identity map. A point $\alpha$ in $\mathbb{C}$ is said to be a **fixed point** of $f$ if $f(\alpha) = \alpha$.
(i) Prove that $f$ has either one or two fixed points.
(ii) Suppose that $f$ has distinct fixed points, $\alpha$ and $\beta$. Prove that

$$\frac{w-\alpha}{w-\beta} = k\frac{z-\alpha}{z-\beta}, \quad \text{where } k = \frac{a-c\alpha}{a-c\beta}.$$

What is the image under $f$ of
(a) the circline $|(z-\alpha)/(z-\beta)| = \lambda$,
(b) the arc $\arg((z-\alpha)(z-\beta)) = \mu \pmod{2\pi}$?
(iii) Suppose that $f$ has a single fixed point $\alpha$. Prove that

$$\frac{1}{w-\alpha} = \frac{1}{z-\alpha} + K, \quad \text{where } K = \frac{1}{a-c\alpha}.$$

2.15 Find the fixed points of the Möbius transformation $z \mapsto w$ when $w$ is given by

$$\text{(i) } \frac{z-1}{z+1}, \quad \text{(ii) } \frac{3z-4}{z-1}, \quad \text{(iii) } \mathrm{i}z, \quad \text{(iv) } \frac{2z-1}{z}.$$

Find the image under each of these mappings of
(a) the circle $|z| = 1$,
(b) the real axis,

(c) the imaginary axis.

2.16 Let $f(z) = 2\mathrm{i}z/(z+\mathrm{i})$. Prove that $f$ maps any circular arc through 0 and i onto itself and deduce that $f$ maps $\{ z : \operatorname{Re} z > 0, \left|z - \frac{1}{2}\mathrm{i}\right| < \frac{1}{2} \}$ to itself. What is the image under $f$ of $\{ z : |z| < |z - \mathrm{i}| \}$?

2.17 (a) Show that the equation $Az\overline{z} + \overline{B}z + B\overline{z} + C = 0$, where $A, C \in \mathbb{R}$ and $B \in \mathbb{C}$, represents a circline. Prove that every circline is representable in this form.

(b) Use (a) to prove that circlines map to circlines under maps of each of the types (rotation, stretching, translation, inversion) considered in 2.13. Hence obtain an alternative proof of Theorem 2.15.

# 3 Topology and analysis in the complex plane

Complex analysis has a vital geometric facet, from which it derives much of its character. The geometry of the plane is inextricably bound up with its topological structure, defined by the open sets we introduce below, and this structure provides the key to the analysis of complex sequences and series and of complex-valued functions defined on subsets of $\mathbb{C}$, in particular the study of convergence, continuity, and differentiability.

## Open sets and closed sets in the complex plane

This section introduces notions which are topological in nature. We shall assume no prior acquaintance with topology and include the bare minimum for our immediate needs. Those who have studied the rudiments of topology in Euclidean spaces or in the more general setting of metric spaces should find much here that is familiar. For an introduction to topology we recommend [**7**].

**3.1 From $\mathbb{R}$ to $\mathbb{C}$.** In elementary real analysis the subsets of $\mathbb{R}$ which principally concern us are intervals. Bounded open intervals, that is, non-empty sets $(c, d) := \{ x \in \mathbb{R} : c < x < d \}$ $(c, d \in \mathbb{R})$ underlie the definitions of limits and continuity; bounded closed intervals (of the form $[c, d] := \{ x \in \mathbb{R} : c \leqslant x \leqslant d \}$ $(c, d \in \mathbb{R})$) are the sets on which continuous functions have especially good behaviour (boundedness, intermediate-value property, etc.). Any bounded open (closed) interval in $\mathbb{R}$ is expressible in the form $\{ x \in \mathbb{R} : |x - a| < r \}$ $(\{ x \in \mathbb{R} : |x - a| \leqslant r \})$; here $a$ is the midpoint and $2r$ the length of the interval. In $\mathbb{C}$, we have, analogously, open discs $\mathrm{D}(a; r)$ $(r > 0)$ and closed discs $\overline{\mathrm{D}}(a; r)$, as defined in 2.6.

**3.2 Definition (open set).** A set $S \subseteq \mathbb{C}$ is **open** if, given $z \in S$, there exists $r > 0$ (depending on $z$) such that $\mathrm{D}(z; r) \subseteq S$. Informally, $S$ is open if, from any point $z$ in $S$, there is room to move some fixed positive distance in any direction without straying outside $S$; how large this distance can be will vary from one point to another (see Fig. 3.1). It is the need for such 'elbow room' that dictates that the sets in so many of our later theorems be open.

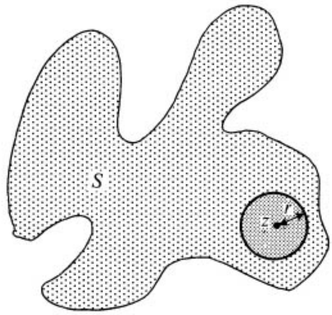



**Figure 3.1** Open set definition

**3.3 Generating open sets.**

- The empty set is open (because the condition for it to be open cannot fail).
- The entire plane $\mathbb{C}$ is open (since $D(z;r) \subseteq \mathbb{C}$ for any $z \in \mathbb{C}$ and any $r > 0$).
- If $S_1, \ldots, S_n$ are open sets, then $S := S_1 \cap \cdots \cap S_n$ is also open. To prove this, let $z \in S$ and pick $\delta_k > 0$ such that $D(z;\delta_k) \subseteq S_k$ for $k = 1, \ldots, n$. Let $\delta := \min\{\delta_1, \ldots, \delta_n\}$. Then $\delta > 0$ (this is where we require the restriction that there are only *finitely many* sets $S_k$). We have $D(z;\delta) \subseteq S$.
- It is easy to prove that, if $S_j$ (for $j \in J$, where $J$ is some index set) are open sets, then $\bigcup_{j \in J} S_j$ is also open.

[The facts here are exactly those required for the family of open subsets of $\mathbb{C}$ to form a **topology** on $\mathbb{C}$.]

**3.4 Examples (open sets).** We return to the examples in 2.6 and show that the sets we called open are indeed open sets. Observe the role played by the *strict* inequalities $<$ and $>$ in these examples.

- We claim that $D(a;r)$ is an open set, for any $a \in \mathbb{C}$ and $r > 0$. To prove this, fix $z \in D(a;r)$ and let $\delta$ satisfy $0 < \delta < r - |z - a|$. Then, using the triangle inequality (1.9(2)),

$$|w - z| < \delta \Longrightarrow |w - a| = |w - z + z - a| \leqslant |w - z| + |z - a| \leqslant \delta + |z - a| < r$$

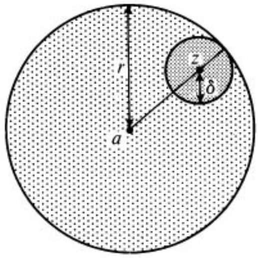



**Figure 3.2** Proving that $\mathrm{D}(a;r)$ is open

(see Fig. 3.2). A similar argument shows that $\{ z \in \mathbb{C} : |z-a| > r \}$ is open.

- The open annulus

$$\{ z \in \mathbb{C} : s < |z-a| < r \} \quad (0 \leqslant s < r < \infty)$$

is the intersection of two open sets and so is itself an open set.

- The open upper half-plane $\Pi^+$ is an open set: if $\operatorname{Im} z > 0$ then $\mathrm{D}(z;r) \subseteq \Pi^+$ whenever $0 < r < \operatorname{Im} z$; More generally, the sector

$$S_{\alpha,\beta} := \{ z \in \mathbb{C} : 0 \neq z = |z|\mathrm{e}^{\mathrm{i}\theta} \in \mathbb{C} \text{ with } \alpha < \theta < \beta \} \quad (\alpha < \beta)$$

is an open set. To prove this, fix $z \in S_{\alpha,\beta}$. Let $\delta := \min\{\delta_1, \delta_2\}$, where $\delta_1$ and $\delta_2$ are the perpendicular distances from $z$ to the bounding rays $\arg z = \alpha$ and $\arg z = \beta$. Then $\delta > 0$ and $\mathrm{D}(z;r) \subseteq S_{\alpha,\beta}$ whenever $0 < r < \delta$.

**3.5 Examples (non-open sets).** The following sets are not open.

- $\overline{\mathrm{D}}(a;r)$: there exists no open disc $\mathrm{D}(z;\delta) \subseteq \mathrm{D}(a;r)$ if $|z-a| = r$; see Fig. 3.3.
- The interval $(0,1)$ of the real axis: for any $z \in (0,1)$ there exists no open disc $\mathrm{D}(z;\delta) \subseteq (0,1)$.

**3.6 Definitions (closed set, limit point, closure).** Let $S \subseteq \mathbb{C}$.

(1) The set $S$ is **closed** if $\mathbb{C} \setminus S = \{ z \in \mathbb{C} : z \notin S \}$ is open.

(2) A point $z \in \mathbb{C}$ is a **limit point** of $S$ if $\mathrm{D}'(z;r) \cap S \neq \varnothing$ for every $r > 0$. A point of $S$ which is not a limit point of $S$ is called an **isolated point** of $S$.

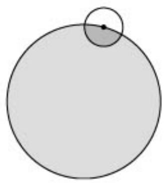


**Figure 3.3** Proving that $\overline{\mathrm{D}}(a;r)$ is not open

(3) The **closure** $\overline{S}$ of $S$ is the union of $S$ and the limit points of $S$.

So $z$ is a limit point of $S$ if every open disc round $z$ contains a point of $S$ other than, possibly, $z$ itself. In Fig. 3.4, $w$ is a limit point of $S$ but $z$ is not. The use of a punctured disc in the definition is to prevent points of $S$ being automatic limit points of $S$. For $z$ to qualify as a limit point it must have clustering round it, arbitrarily closely, points of $S$ (other than $z$ itself if $z \in S$).

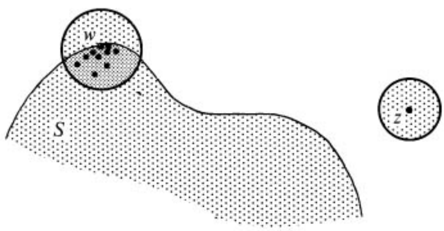



**Figure 3.4** Limit points

The definitions in 3.6 are linked by the following proposition. It is useful to put this in place before we consider examples.

**3.7 Proposition (closed sets and closure).** Let $S \subseteq \mathbb{C}$.

(1) The following are equivalent:

(i) $S$ is closed,

(ii) $S$ contains all its limit points,

(iii) $\overline{S} = S$.

(2) $z \in \overline{S}$ if and only if $V \cap S \neq \varnothing$ for every open set $V$ containing $z$.

(3) $\overline{S}$ is a closed set.

**Proof** (1) Note that we have $\mathrm{D}(z;r) \cap S = \mathrm{D}'(z;r) \cap S$ for $z \notin S$. Then

$$\begin{aligned} &S \text{ is closed} \\ &\quad \iff \mathbb{C} \setminus S \text{ is open} \\ &\quad \iff \text{given } z \notin S, \text{ there exists } r > 0 \text{ such that } \mathrm{D}(z;r) \subseteq \mathbb{C} \setminus S \\ &\quad \iff \text{given } z \notin S, \text{ there exists } r > 0 \text{ such that } \mathrm{D}'(z;r) \cap S = \varnothing \\ &\quad \iff \text{no point of } \mathbb{C} \setminus S \text{ is a limit point of } S. \end{aligned}$$

Hence (i) and (ii) are equivalent. Also, (ii) and (iii) are equivalent by definition of $\overline{S}$.

Now consider (2). Suppose that every open set $V$ containing $z$ intersects $S$. Recall from 3.4 that $\mathrm{D}(z;r)$ is an open set containing $z$ for any $r > 0$. Hence either $z \in S$ or $\mathrm{D}(z;r) \cap S \neq \varnothing$ (for all $r > 0$). For the converse, suppose that there is some open set $V$ such that $z \in V$ and $V \cap S = \varnothing$. Since $V$ is open, we can choose $r > 0$ such that $\mathrm{D}(z;r) \subseteq V$. Then $\mathrm{D}(z;r) \cap S = \varnothing$.

For (3), it suffices by (1) to prove that any $z$ in $\overline{\overline{S}}$ is in $\overline{S}$. Suppose, for a contradiction, that this is false. Because $z \notin \overline{S}$, there exists $r > 0$ such that $\mathrm{D}(z;r) \cap S = \varnothing$. Because $z \in \overline{\overline{S}}$, there exists $w$ such that $w \in \mathrm{D}(z;r) \cap \overline{S}$ (by (2), applied with $\overline{S}$ in place of $S$). But then $\mathrm{D}(z;r)$ is an open set containing $w$ and, by (2) again, we have $\mathrm{D}(z;r) \cap S \neq \varnothing$, the required contradiction. □

### 3.8 Examples (closed sets, limit points, closure).

- Consider $\overline{\mathrm{D}}(a;r)$. We saw in 3.4 that the complement of this set is open, so $\overline{\mathrm{D}}(a;r)$ is closed. So, too, is the circle $\{ z \in \mathbb{C} : |z - a| = r \}$.
  A point $z$ is a limit point of $\mathrm{D}(a;r)$ if and only if $|z - a| \leqslant r$. Hence the closure of $\mathrm{D}(a;r)$ is the closed disc $\overline{\mathrm{D}}(a;r)$.
- Half-planes, sectors, annuli, etc., when specified by weak inequalities ($\leqslant$) are closed, and are the closures of the corresponding sets specified by strict or mixed inequalities.
- The set $S = \{ z : 1 \leqslant |z| < 2 \}$ is neither open nor closed. To see this, note that no disc $\mathrm{D}(1;r) \subseteq S$ and that no disc $\mathrm{D}(2;r) \subseteq \mathbb{C} \setminus S$. Sets in $\mathbb{C}$ are not like doors: a set which fails to be open need not be a closed set. The concepts open and closed are related by set complementation, not by logical negation.
- Let $S = \{ (-1)^n(1 + \frac{1}{n}) : n = 1, 2, \ldots \}$. Certainly $S$ is not open. The limit points of $S$ are $\pm 1$, and neither of these belongs to $S$. Hence $S$ is not closed.

**3.9 Bounded sets and compact sets.** A subset $S$ of $\mathbb{C}$ is said to be **bounded** if there exists a real constant $M$ such that $|z| \leqslant M$ for all $z \in S$. A set which is both bounded and closed is called **compact**. [The sets we have called compact are exactly those which satisfy the usual open covering definition of compactness, thanks to the Heine–Borel Theorem. Save in Theorem 4.5, where we proceed *ad hoc*, we shall not need to work with open coverings.]

Note in particular that line segments $[\alpha, \beta]$ in $\mathbb{C}$, circles $|z - a| = r$, and closed discs $\overline{\mathrm{D}}(a; r)$ are compact.

**3.10 The extended plane revisited.** We defined open sets in $\mathbb{C}$ via open discs. This suggests that we should define 'discs' in $\widetilde{\mathbb{C}}$ centred on $\infty$: we let

$$\mathrm{D}(\infty; r) := \{\, z \in \mathbb{C} : |z| > r \,\} \cup \{\infty\} \quad (r > 0).$$

We define a subset $S$ of $\widetilde{\mathbb{C}}$ to be open if, for each $z \in S$, there exists $r > 0$ such that $\mathrm{D}(z; r) \subseteq S$. The motivation here comes from stereographic projection. [Those with the requisite topological knowledge will see that the bijective correspondence between the Riemann sphere $\Sigma$ and the extended plane $\widetilde{\mathbb{C}}$ sets up a homeomorphism between $\Sigma$, equipped with the topology it acquires as a subspace of the Euclidean space $\mathbb{R}^3$, and $\widetilde{\mathbb{C}}$. Furthermore, $\Sigma$ is compact and, since compactness is preserved by homeomorphism, $\widetilde{\mathbb{C}}$ is also a compact topological space. What we have here is an instance of the 'one-point compactification' of a space: by adding the extra point $\infty$ to $\mathbb{C}$ and defining open sets appropriately in $\widetilde{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ we obtain a compact space.]

## Convexity and connectedness

This section deals with concepts with both geometric and topological content. The notion of a region (defined in 3.12) is the most important for the future development of the theory.

**3.11 Convex sets and polygonally connected sets.** Let $S \subseteq \mathbb{C}$. We say that $S$ is **convex** if, given any pair of points $a$ and $b$ in $S$, we have $[a, b] \subseteq S$. See Fig. 3.5. Examples of convex sets are:

- half-planes and, more generally, sectors;
- open discs $\mathrm{D}(a; r)$ and closed discs $\overline{\mathrm{D}}(a; r)$.

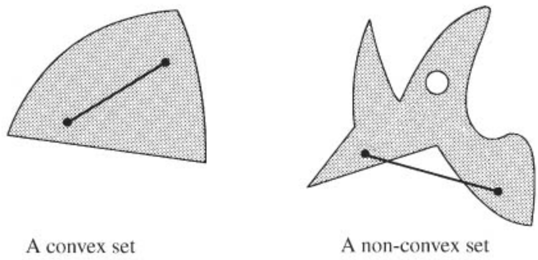



**Figure 3.5** Convexity

Examples of non-convex sets are:

- the union of non-intersecting discs, for example $\mathrm{D}(-1;1) \cup \mathrm{D}(1;1)$;
- the union of two intersecting discs, when neither lies inside the other;
- $\mathbb{C} \setminus \mathbb{R}$, the plane with the real axis removed;
- $\mathbb{C} \setminus [0,\infty)$, the plane with the non-negative real axis removed;
- punctured discs $\mathrm{D}'(a;r)$ or, more generally, annuli.

The non-convexity of $\mathbb{C} \setminus \mathbb{R}$ seems geometrically obvious but, for reasons which will emerge later, it is worth verifying this analytically too. Let $a = a_1 + \mathrm{i}a_2$ and $b = b_1 + \mathrm{i}b_2$ with $a_2 < 0$ and $b_2 > 0$. A typical point of the line segment $[a,b]$ is

$$\begin{aligned}(1-t)a + tb &= (1-t)a_1 + \mathrm{i}(1-t)a_2 + (tb_1 + \mathrm{i}tb_2) \\ &= ((1-t)a_1 + tb_1) + \mathrm{i}((1-t)a_2 + tb_2) \quad (0 \leqslant t \leqslant 1).\end{aligned}$$

But $t = a_2(a_2 - b_2)^{-1} \in [0,1]$ and yet for this $t$ we have $\mathrm{Im}((1-t)a + tb) = 0$. So $[a,b] \not\subseteq \mathbb{C} \setminus \mathbb{R}$.

There is an important difference, geometrically, between $S_1 := \mathbb{C} \setminus \mathbb{R} = \Pi^+ \cup \Pi^-$ and $S_2 := \mathbb{C} \setminus [0,\infty)$. First consider $S_2$ and fix any two points $a, b \in S_2$. Although we may not be able to join these points by a *single* line segment lying wholly in $S_2$, we can 'walk round the slit': there exist points $z_0 = a = a_1 + \mathrm{i}a_2$, $z_1 = -1 + \mathrm{i}a_2$, $z_2 = -1 + \mathrm{i}b_2$, $z_3 = b = b_1 + \mathrm{i}b_2$ such that the union of the line segments $[z_0, z_1]$, $[z_1, z_2]$, and $[z_2, z_3]$ lies entirely within $S_2$. On the other hand, given points $a \in \Pi^+$ and $b \in \Pi^-$, there is no finite sequence of line segments $[z_0, z_1], [z_1, z_2], \ldots, [z_{n-1}, z_n]$ with $z_0 = a$ and $z_n = b$ which lies wholly in $S_1$. (For a proof, observe that if there were such a sequence then there would be some $k$ such that $[z_{k-1}, z_k] \subseteq S_1$, which is impossible, from above.)

Let $a = z_0, z_1, \ldots, z_{n-1}, b = z_n$ be finitely many points in $\mathbb{C}$. We call

$$[z_0, z_1] \cup [z_1, z_2] \cup \cdots \cup [z_{n-1}, z_n]$$

a **polygonal route from** $a$ **to** $b$; the case $a = b$, with $n = 0$, is allowed. A non-empty subset $S$ of $\mathbb{C}$ is **polygonally connected** if, given any two points $a$ and $b$ in $S$, there is a polygonal route from $a$ to $b$ lying wholly in $S$. Clearly every convex set is polygonally connected. Examples of sets which are polygonally connected but not convex are the set $S_1$ defined above, and any annulus.

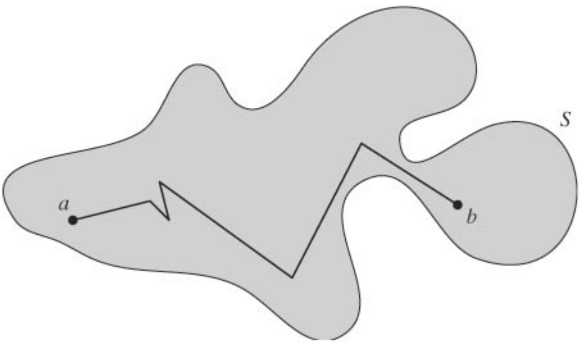



**Figure 3.6** A polygonal route from $a$ to $b$ in $S$

Polygonal connectedness is related to a condition expressible in terms of open sets, as Theorem 3.13 will show.

**3.12 Definitions (connectedness, region).**

(1) A subset $G$ of $\mathbb{C}$ is **connected** if it cannot be expressed as the union of non-empty open sets $G_1$ and $G_2$ with $G_1 \cap G_2 = \varnothing$. Putting this another way, connectedness of $G$ implies that if $G_1 \subseteq G$ and $G_1$ and $G \setminus G_1$ are both open, then we must have $G_1 = G$ or $G_1 = \varnothing$.

(2) A **region** is a non-empty open connected subset of $\mathbb{C}$.

**3.13 Theorem.** Let $G$ be a non-empty open subset of $\mathbb{C}$. Then $G$ is a region if and only if $G$ is polygonally connected. In particular, any non-empty open convex set is a region.

**Proof** Suppose that $G$ is a region. Fix $a \in G$ and let

$$G_1 := \{ z \in G : \exists \text{ a polygonal route from } a \text{ to } z \text{ in } G \}$$

and $G_2 := G \setminus G_1$. Certainly $a \in G_1$, so $G_1 \neq \varnothing$. Our strategy will be to show that both $G_1$ and $G_2$ are open. Then connectedness implies that $G = G_1$.

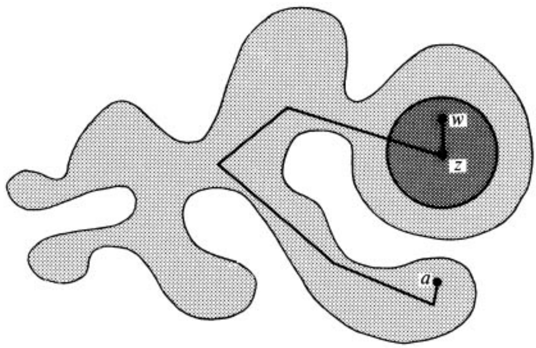



**Figure 3.7** Proof of Theorem 3.13

We now establish our claim that $G_1$ and $G_2$ are open. For any $z \in G$ we can find $r > 0$ such that $\mathrm{D}(z;r) \subseteq G$. Let $w \in \mathrm{D}(z;r)$. Certainly $[z,w] \subseteq \mathrm{D}(z;r) \subseteq G$. If $z \in G_1$ then there is a polygonal route in $G$ from $a$ to $w$, via $z$, so that $w \in G_1$. On the other hand, if $w \in G_1$ then there is a polygonal route in $G$ from $a$ to $z$, via $w$. Hence if $z \notin G_1$ then $w \notin G_1$. We conclude that $z \in G_k$ implies $\mathrm{D}(z;r) \subseteq G_k$ for $k = 1, 2$. Therefore $G_1$ and $G_2$ are open, as claimed.

Conversely, suppose that $G$ is non-empty, open, and polygonally connected, and suppose for a contradiction that $G$ is the disjoint union of non-empty open sets $G_1$ and $G_2$. Take any $a \in G_1$ and $b \in G_2$, and a polygonal route $[z_0, z_1] \cup \cdots \cup [z_{n-1}, z_n]$ in $G$ which joins $a = z_0$ to $b = z_n$. Then at least one of the line segments $[z_{k-1}, z_k]$ has $z_{k-1} \in G_1$ and $z_k \in G_2$. A typical point of $[z_{k-1}, z_k]$ is $z(t) := (1-t)z_{k-1} + tz_k$ with $0 \leqslant t \leqslant 1$ and we have the following:

(i) for each $t \in [0,1]$ either $z(t)$ lies in $G_1$ or it lies in $G_2$, but not both;

(ii) $z(0) = z_{k-1} \in G_1$ and $z(1) = z_k \in G_2$;

(iii) if $z(t)$ lies in $G_k$ for some given $t$, then $z(s) \in G_k$ for all $s \in [0,1]$ with $|t-s|$ sufficiently small (because $G_k$ is open).

Let

$$q := \sup\{ t \in [0,1] : z(t) \in G_1 \}.$$

By (i) and (iii), we have $0 < q < 1$. Now consider $z(q)$. If $z(q) \in G_1$ then, by (iii), there exists $s$ with $1 > s > q$ such that $z(s) \in G_1$, in contradiction to the definition of $q$. If $z(q) \in G_2$ then we can find $\delta > 0$ such that $z(s) \in G_2$ (and hence $z(s) \notin G_1$) for all $s$ such that $0 < q - \delta < s \leqslant q$. Again we have a contradiction to the definition of $q$. □

[Those with prior knowledge of connectedness will have recognized that the argument above has features in common with that used to prove that an interval in $\mathbb{R}$ is connected.]

**3.14 Other characterizations of polygonal connectedness.** In our definition of polygonal connectedness we used polygonal routes made up of arbitrary line segments. There are variants on Theorem 3.13. Essentially the same proof works if, for example,

- we allow only polygonal routes all of whose line segments are either horizontal or vertical, or
- we allow circular arcs in place of some or all of the line segements in a polygonal route.

## Limits and continuity

This section contains the technical foundations of analysis in the complex plane. It deals only with those concepts and results which transfer in an entirely straightforward manner from elementary real analysis. We assume that the reader is already familiar with limits and continuity in the real case, as presented in introductory analysis texts. We give references to [**3**], which covers all the real analysis we require here and in succeeding chapters, but any other comparable text serves equally well as a source.

The definitions which follow mimic those for real sequences and functions, with open discs in $\mathbb{C}$ replacing open intervals in $\mathbb{R}$.

**3.15 Definitions (sequences).**

(1) A **(complex) sequence** $\{z_n\}$ is an assignment of a complex number $z_n$ to each natural number $n = 1, 2, \ldots$. We occasionally need to allow a different set of values of $n$; notation such as $\{z_n\}_{n \geqslant 0}$ should be self-explanatory. The sequence $\{z_n\}$ is **bounded** if there is a real constant $M$ such that $|z_n| \leqslant M$ for all $n$.

(2) The sequence $\{z_n\}$ **converges, to the limit** $a$ in $\mathbb{C}$ (in symbols, $z_n \to a$ or $\lim_{n\to\infty} z_n = a$), if, given $\varepsilon > 0$, there exists a natural number $N$ (depending on $\varepsilon$) such that

$$n \geqslant N \quad \Longrightarrow \quad |z_n - a| < \varepsilon.$$

(3) The sequence $\{w_k\}$ is a **subsequence** of the sequence $\{z_n\}$ if there exist natural numbers $n_1 < n_2 < \ldots$ such that $w_k = z_{n_k}$ for $k = 1, 2, \ldots$.

**3.16 Definitions (limits of functions, continuity).**

(1) Let $f\colon S \to \mathbb{C}$ be a function defined on a set $S \subseteq \mathbb{C}$ and let $a \in \overline{S}$. Then $\lim_{z\to a} f(z) = w$ (in alternative notation, $f(z) \to w$ as $z \to a$) if, given $\varepsilon > 0$, there exists $\delta > 0$ (depending on $a$ and $\varepsilon$) such that

$$z \in S \text{ and } 0 < |z - a| < \delta \quad \Longrightarrow \quad |f(z) - w| < \varepsilon.$$

Note the inclusion, as is normal with limits of functions, of the condition $0 < |z - a|$, that is, $z \neq a$. The limit, if it exists, is determined by the behaviour of $f(z)$ as $z$ *approaches* $a$; the value of $f$ *at* $a$ is irrevelant, and may not even be defined if $a \notin S$.

(2) Let $f\colon S \to \mathbb{C}$ be a function. Then $f$ is **continuous at** $a \in S$ if, given $\varepsilon > 0$, there exists $\delta > 0$ (depending on $a$ and $\varepsilon$) such that

$$z \in S \text{ and } |z - a| < \delta \quad \Longrightarrow \quad |f(z) - f(a)| < \varepsilon.$$

In other words, $\lim_{z\to a} f(z)$ exists and equals $f(a)$.
The function $f$ is **continuous on** $S$ if it is continuous at each $a \in S$. [Continuous functions can be more elegantly characterized in terms of open sets, but we shall require only the $\varepsilon$-$\delta$ definition.]

**3.17 Working with limits.** The algebra of complex limits (sums, products, etc.) and other elementary properties can be developed, both for sequences and for functions, exactly as in the real case. We shall assume this done, and shall use the results freely. One caveat is, however, necessary. Proofs which depend on the order structure of $\mathbb{R}$ do not transfer directly to $\mathbb{C}$. It is therefore useful to have available the following elementary result.

**3.18 Lemma (linking convergence in $\mathbb{C}$ and in $\mathbb{R}$).**

(1) Let $\{z_n\}$ be a complex sequence. Then $\{z_n\}$ converges if and only if the real sequences $\{\operatorname{Re} z_n\}$ and $\{\operatorname{Im} z_n\}$ both converge. In addition, if $z_n \to a$ then $|z_n| \to |a|$ and $\overline{z_n} \to \overline{a}$.

(2) Let $f\colon S \to \mathbb{C}$ and write, as usual, $f = u + \mathrm{i}v$, where $u$ and $v$ are real-valued. To parallel (1), write $u$ as $\operatorname{Re} f$ and $v$ as $\operatorname{Im} f$, so

$$(\operatorname{Re} f)(z) := \operatorname{Re}(f(z)) \quad \text{and} \quad (\operatorname{Im} f)(z) := \operatorname{Im}(f(z)) \quad (z \in S).$$

Then, for any $a \in \overline{S}$, $\lim_{z\to a} f(z)$ exists if and only if both $\lim_{z\to a} \operatorname{Re} f(z)$ and $\lim_{z\to a} \operatorname{Im} f(z)$ exist. Then $f(z) \to w$ implies that $\operatorname{Re} f(z) \to \operatorname{Re} w$ and $\operatorname{Im} f(z) \to \operatorname{Im} w$ and, in addition, $|f(z)| \to |w|$ and $\overline{f(z)} \to \overline{w}$.

(3) Let $f\colon S \to \mathbb{C}$. Then $f$ is continuous at $a \in S$ (on $S$) if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are continuous at $a$ (on $S$). In addition, continuity of $f$ implies continuity of $|f|$, defined by $|f|(z) := |f(z)|$ $(z \in S)$.

**Proof** We shall prove (1), leaving as an exercise the proofs of the analogous statements in (2) and (3). Let $z_n \to a$. Using 1.8(1),

$$\begin{aligned} |\operatorname{Re} z_n - \operatorname{Re} a| &= |\operatorname{Re}(z_n - a)| \leqslant |z_n - a| \quad \text{and} \\ |\operatorname{Im} z_n - \operatorname{Im} a| &= |\operatorname{Im}(z_n - a)| \leqslant |z_n - a| . \end{aligned}$$

It then follows easily from the definition of convergence that $\operatorname{Re} z_n \to \operatorname{Re} a$ and $\operatorname{Im} z_n \to \operatorname{Im} a$. Conversely, if $\operatorname{Re} z_n \to p$ and $\operatorname{Im} z_n \to q$, then

$$z_n = \operatorname{Re} z_n + \mathrm{i} \operatorname{Im} z_n \to p + \mathrm{i}q,$$

by the complex version of the algebra of limits. Also, from above, $\operatorname{Re} z_n \to p$ and $\operatorname{Im} z_n \to q$. By uniqueness of real limits, $\operatorname{Re} a = p$ and $\operatorname{Im} a = q$.

To prove the claim about $\{|z_n|\}$ we observe that, for any $a \in \mathbb{C}$, 1.8(3) gives

$$||z_n| - |a|| \leqslant |z_n - a| .$$

Finally, if $z_n \to a = p + \mathrm{i}q$ then $\overline{z_n} = \operatorname{Re} z_n - \mathrm{i} \operatorname{Im} z_n \to p - \mathrm{i}q = \overline{a}$. $\square$

**3.19 Examples (limits of sequences).**

- Consider $\{z_n\}$ where $z_n = c^n$, for $c \in \mathbb{C}$ and $|c| \neq 1$. If $|c| < 1$ then $|z_n| = |c|^n \to 0$. On the other hand, if $|c| > 1$ then the real sequence $\{|c|^n\}$ tends to infinity and hence $\{z_n\}$ has no limit in $\mathbb{C}$.

- Now consider $\{z_n\}$ where $z_n = (n+\mathrm{i})^{-1}$. Here we may note that

$$z_n = \frac{n}{n^2+1} - \frac{1}{n^2+1}\mathrm{i}$$

and deduce that $z_n \to 0$ since

$$\operatorname{Re} z_n = n/(n^2+1) \to 0 \quad \text{and} \quad \operatorname{Im} z_n = -1/(n^2+1) \to 0.$$

Alternatively, we see that $|z_n| \leqslant (n-1)^{-1}$ for $n \geqslant 2$, by 1.8(3). This implies that $|z_n| \to 0$ and hence that $z_n \to 0$.

**3.20 Examples (limits of functions, continuity).**

- Consider $f(z) = (\operatorname{Im} z)/(\operatorname{Re} z)$ for $z \neq 0$. Then, in particular, $f(z) = 0$ when $z$ is real and $f(z) = 1$ on the line $y = x$. Hence $\lim_{z\to 0} f(z)$ fails to exist.
- $z^{100} - 1$ is continuous on $\mathbb{C}$ and $(z^{100}-1)^{-1}$ is continuous except at the 100th roots of unity.
- From 3.18(2), we see that $\operatorname{Re} z$, $\operatorname{Im} z$, and $\overline{z}$ are continuous functions of $z$ on $\mathbb{C}$.

The following important result is almost immediate from the corresponding result about $\mathbb{R}$ and the lemma in 3.18.

**3.21 The Cauchy convergence principle for $\mathbb{C}$.** A complex sequence $\{z_n\}$ converges if and only if, given $\varepsilon > 0$, there exists a natural number $N$ such that

$$m, n \geqslant N \Longrightarrow |z_m - z_n| < \varepsilon.$$

We next use 3.18(1) to derive the following theorem from its real counterpart (given, for example, in **[3]**, 3.4.8). This theorem leads to two further results which are needed in the course of proving some important theorems later on.

**3.22 Convergent subsequence theorem.** Any bounded sequence of complex numbers has a convergent subsequence.

**Proof** Let $\{z_n\}$ be a sequence with $|z_n| \leqslant M$ for all $n$. Then, by 1.9(1), $|\operatorname{Re} z_n| \leqslant M$, so $\{\operatorname{Re} z_n\}$ is a bounded sequence in $\mathbb{R}$. Hence there exist natural numbers $n_1 < n_2 < \ldots$ such that $\{\operatorname{Re} z_{n_k}\}_{k\geqslant 1}$ converges. By 1.9(1) again, $\{\operatorname{Im} z_{n_k}\}_{k\geqslant 1}$ is a bounded real sequence, so there exist natural numbers $m_j = n_{k_j}$

with $m_1 < m_2 < \dots$ such that $\{\operatorname{Im} z_{m_j}\}_{j\geqslant 1}$ converges. Also $\{\operatorname{Re} z_{m_j}\}_{j\geqslant 1}$ converges because it is a subsequence of $\{\operatorname{Re} z_{n_k}\}_{k\geqslant 1}$. Now 3.18(1) shows that $\{z_{m_j}\}_{j\geqslant 1}$ provides a convergent subsequence of $\{z_n\}$. □

**3.23 Corollary (Bolzano–Weierstrass theorem).** Any infinite compact subset $S$ of $\mathbb{C}$ has a limit point in $S$.

**Proof** We first observe that a point $z$ is a limit point of $S$ if and only if there exists a sequence $\{w_k\}_{k\geqslant 1}$ of distinct points of $S$ such that $w_k \to z$. We leave the proof as an exercise.

By the definition in 3.9, $S$ is bounded and closed. Select a sequence $\{z_n\}$ with the points $z_n$ distinct and belonging to $S$. Theorem 3.22 asserts that $\{z_n\}$ has a subsequence which converges, to $z$ say. Then $z$ is a limit point of $S$. Because $S$ is closed, it contains $z$ (by Proposition 3.7). □

**3.24 Boundedness theorem for continuous functions.** Let $S$ be a compact subset of $\mathbb{C}$ and $f\colon S \to \mathbb{C}$ a continuous function. Then

(1) $f$ is bounded, that is, there exists a finite constant $M$ such that $|f(z)| \leqslant M$ for all $z \in S$;

(2) $f$ attains its bounds, that is, there exist $z_1$ and $z_2$ in $S$ such that

$$|f(z_1)| \leqslant |f(z)| \leqslant |f(z_2)| \quad \text{for all } z \in S.$$

We record for future use one further theorem from real analysis. It is an immediate corollary of the Intermediate value theorem ([**3**], 5.3.7).

**3.25 Theorem (integer-valued continuous functions).** Let $[a, b]$ be a closed bounded subinterval in $\mathbb{R}$ and let $f\colon [a, b] \to \mathbb{Z}$ be continuous. Then $f$ is constant.

## Exercises

**Exercises from the text.** Prove that the union of any family of open subsets of $\mathbb{C}$ is open (see 3.3). Verify the claims in 3.11 and the unproved assertions in 3.18. Prove the statement about limit points in the first paragraph of the proof in 3.23.

3.1 (a) Prove that the following are open sets:

$$\text{(i) } \{ z \in \mathbb{C} : |z-1| < |z+\mathrm{i}| \}, \qquad \text{(ii) } \mathbb{C} \setminus [0,1].$$

(b) Prove that the following are not open:

$$\text{(i) } \{ z \in \mathbb{C} : \operatorname{Re} z \geqslant 0 \}, \qquad \text{(ii) } \{ z \in \mathbb{C} : |z| \leqslant 2, \operatorname{Re} z > 1 \}.$$

3.2 Which of the sets in Exercise 2.2 are open, which are closed, and which are neither? Find the closures of the non-closed sets.

3.3 Prove that each of the sets in Exercise 2.3 is open. Which are convex? Which are regions?

3.4 Let $S$ be a finite subset of $\mathbb{C}$.
(i) Prove that $S$ is open only if $S = \varnothing$.
(ii) Prove that $S$ is closed.

3.5 (a) Let $G_1$ and $G_2$ be regions in $\mathbb{C}$. Prove that $G_1 \cup G_2$ is a region if and only if $G_1 \cap G_2 \neq \varnothing$.
(b) Let $G_1, \ldots, G_N$ be regions in $\mathbb{C}$ such that $G_k \cap G_{k+1} \neq \varnothing$ for $k = 1, \ldots, N-1$. Prove that $\bigcup_{k=1}^{N} G_k$ is a region.

3.6 Suppose that $G$ is a region and let $a \in G$. Prove that $G \setminus \{a\}$ is a region. Assume that $r$ is such that $\mathrm{D}(a;r) \subseteq G$. Is $G \setminus \mathrm{D}(a;r)$ always a region?

3.7 Let $S$ be a subset of $\mathbb{C}$ and let $T := \{ \overline{z} : z \in S \}$. Prove that $S$ is open (a region) if and only if $T$ is open (a region).

3.8 Let $G$ be a region and $f$ a continuous map of $G$ onto an open set $\widehat{G}$. Prove that $\widehat{G}$ is a region.

3.9 (a) Prove that the sequence $\{z_n\}$ converges, and give its limit, when $z_n$ is given by

$$\text{(i) } \frac{1}{n}\mathrm{i}^n, \quad \text{(ii) } (1+\mathrm{i})^{-n}, \quad \text{(iii) } \frac{n^2+\mathrm{i}n}{n^2+\mathrm{i}}.$$

(b) Prove that the sequence $\{z_n\}$ does not converge when $z_n$ is given by

$$\text{(i) } \mathrm{i}^n, \quad \text{(ii) } (1+\mathrm{i})^n, \quad \text{(iii) } (-1)^n\frac{n}{n+\mathrm{i}}.$$

3.10 Let $p$ be any complex number. Let $z_0 = p$ and, for $n \geqslant 1$, define

$$z_{n+1} = \frac{1}{2}\Big(z_n - \frac{1}{z_n}\Big),$$

if $z_n \neq 0$. Prove the following assertions.

(i) If $\{z_n\}$ converges to a limit $a$ then $a^2 + 1 - 0$.

(ii) If $p$ is real, then $\{z_n\}$, if defined, does not converge.

(iii) If $p = \mathrm{i}q$, where $q \in \mathbb{R} \setminus \{0\}$, then $\{z_n\}$ converges.

(iv) If $|p| = 1$ and $p \neq \pm 1$, then $\{z_n\}$ converges.

(v) If $\operatorname{Im} p > 0$, then $\{z_n\}$ converges to $1$ and if $\operatorname{Im} p < 0$ then $\{z_n\}$ converges to $-\mathrm{i}$ (hint: consider $|z_n - \mathrm{i}| \,/\, |z_n + \mathrm{i}|$).

3.11 (a) Prove that $\lim_{z \to 0} f(z)$ exists and equals 0 for each of the following choices of $f(z)$:

$$\text{(i) } z + |z|^3, \quad \text{(ii) } \frac{|z|^2}{z}, \quad \text{(iii) } \frac{(\operatorname{Re} z)(\operatorname{Im} z)}{|z|}.$$

(b) For each of the following choices of $f$ prove, by letting $z$ approach 0 along suitable rays, that $\lim_{z \to 0} f(z)$ fails to exist:

$$\text{(i) } \frac{\bar{z}}{z}, \quad \text{(ii) } \frac{\bar{z}}{|z|}, \quad \text{(iii) } \frac{\operatorname{Im} z}{\operatorname{Re} z}.$$

3.12 On which sets in $\mathbb{C}$ are the following continuous?

$$\text{(i) } (z^2 - 1)^{-1}, \quad \text{(ii) } \frac{z - \mathrm{i}}{2z}, \quad \text{(iii) } \frac{z^2 + 1}{z^3 + 1}, \quad \text{(iv) } \frac{z}{(\operatorname{Re} z)^2}.$$

3.13 Define a function $f$ by $f(z) = z/(1 + |z|)$.

(i) Prove that $f$ is continuous on $\mathbb{C}$.

(ii) Prove that $f(z_1) = f(z_2)$ implies $z_1 = z_2$ (hint: use polar coordinates).

(iii) Prove that $f$ maps $\mathbb{C}$ onto $\mathrm{D}(0;1)$ (hint: use polar coordinates).

(Hence $f$ is a continuous bijection from $\mathbb{C}$ onto $\mathrm{D}(0;1)$.)

3.14 Suppose that $S$ is a compact subset of $\mathbb{C}$ and $f\colon S \to \mathbb{C}$ a continuous function. Use Theorem 3.24 to prove that the image $f(S)$ of $S$ under $f$ is compact.

3.15 (The result of this exercise is needed in 17.9; for those who know some topology it is a direct consequence of the Bolazano–Weierstrass theorem for $\Sigma \subseteq \mathbb{R}^3$.) Let $S$ be an infinite subset of $\widetilde{\mathbb{C}}$. Show that either $S$ has a limit point in $S$ or every disc $\mathrm{D}'(\infty; r)$ contains a point of $S$ [so that $\infty$ is a limit point of $S$ in the space $\widetilde{\mathbb{C}}$].

3.16 Let $c \in \mathbb{C}$ and let $z_n = z_n(c)$ (depending on $c$) be defined inductively as follows:

$$z_0 = 0, \quad z_{n+1} = F_c(z_n) \ \ (\text{for } n \geqslant 0), \quad \text{where } F_c(z) := z^2 + c.$$

The **Mandelbrot set** $M$ can be defined to be the set of $c \in \mathbb{C}$ for which $|c| \leqslant 2$ and the sequence $\{|z_n(c)|\}$ does not tend to infinity as $n \to \infty$.

(i) Suppose there exists $k$ such that $|z_k(c)| > 2$. Prove that $c \notin M$.

(ii) Let $U_k := \{\, c \in \mathbb{C} : |z_k(c)| > 2 \,\}$. Prove that $U_k$ is open.

(iii) Deduce that $M$ is a closed subset of $\mathbb{C}$.

(The Mandelbrot set is discussed in the Appendix.)

# 4 Paths

In this chapter we start systematically to study curves in the complex plane. To regard a curve, such as a circle, simply as a subset of the plane will not suffice for our purposes. Instead, we adopt a dynamic approach and think of a curve as the route traced out by a moving point, the route being specified by a suitable function of some real parameter. For example, $\gamma(t) = e^{it}$ travels anticlockwise once round the unit circle as $t$ increases from $0$ to $2\pi$, starting from $1$.

We shall need a few of the ideas from this chapter when we consider elementary conformal mapping in Chapter 8 and multifunctions in Chapter 9. Readers seeking the quickest possible route to the theory at the heart of complex analysis, in particular to Cauchy's theorem, may wish to proceed straight from this chapter to Chapter 10, in which we discuss integrals along paths.

## Introducing curves and paths

Since the dynamic approach to curves involves a function $\gamma$ of a real variable $t$ which takes values in $\mathbb{C}$, we record a few facts about such functions.

**4.1 Complex-valued functions defined on real intervals.** Suppose $g\colon [\alpha, \beta] \to \mathbb{C}$ is a function. Then $g$ is continuous if and only if its real and imaginary parts $\operatorname{Re} g$ and $\operatorname{Im} g$ are continuous (recall 3.16 and 3.18). Differentiability of $g$ is defined via the expected limit:

$$g'(t) := \lim_{h \to 0} \frac{g(t+h) - g(t)}{h} \qquad (t, t+h \in [\alpha, \beta]),$$

if this limit exists. Derivatives at $\alpha$ and $\beta$ are, of course, one-sided derivatives. Using 3.18(2), we see that $g'(t)$ exists if and only if $(\operatorname{Re} g)'(t)$ and $(\operatorname{Im} g)'(t)$ exist, and then $g'(t) = (\operatorname{Re} g)'(t) + i(\operatorname{Im} g)'(t)$.

We now present the formal definitions of curves of various types. These are illustrated in Fig. 4.1.

**4.2 Curves and paths.** Let $[\alpha,\beta]$ $(-\infty < \alpha \leqslant \beta < \infty)$ be a closed bounded interval in $\mathbb{R}$. A **curve** $\gamma$ **with parameter interval** $[\alpha,\beta]$ is a continuous function $\gamma\colon [\alpha,\beta] \to \mathbb{C}$. It has **initial point** $\gamma(\alpha)$ and **final point** $\gamma(\beta)$, and is **closed** if $\gamma(\alpha) = \gamma(\beta)$. It is **simple** if $\alpha \leqslant s < t \leqslant \beta$ implies that $\gamma(s) \neq \gamma(t)$ unless $s = t$ or, in the case of a closed curve, $s = \alpha$ and $t = \beta$.

Suppose $\gamma$ is a curve with parameter interval $[\alpha,\beta]$. We denote the image

$$\gamma([\alpha,\beta]) := \{\, \gamma(t) \,:\, t \in [\alpha,\beta] \,\}$$

by $\gamma^*$. The curve $\gamma$ is said to **lie in a set** $S$ if $\gamma^* \subseteq S$. As the continuous image of a compact interval, $\gamma^*$ is a compact subset of $\mathbb{C}$ (by Exercise 3.14). In particular, $\gamma^*$ is a closed set.

The curve $\gamma$ carries a built-in orientation, determined by the direction in which $\gamma(t)$ traces out $\gamma^*$ as $t$ increases from $\alpha$ to $\beta$. Given $\gamma$, there exists a curve $-\gamma$ with the same image set but the opposite orientation:

$$(-\gamma)(t) := \gamma(\alpha + \beta - t) \quad (t \in [\alpha,\beta]).$$

Let $\alpha \leqslant \alpha_1 \leqslant \beta_1 \leqslant \beta$. By restricting the function $\gamma$ to $[\alpha_1,\beta_1]$, we obtain a new curve, which we denote by $\gamma\restriction[\alpha_1,\beta_1]$. Now suppose $\alpha < \tau < \beta$ and let $\gamma_1 = \gamma\restriction[\alpha,\tau]$ and $\gamma_2 = \gamma\restriction[\tau,\beta]$. The final point of $\gamma_1$ coincides with the initial point of $\gamma_2$ (each is $\gamma(\tau)$) and $\gamma^*$ is traced by first tracing $\gamma_1^*$ and then tracing $\gamma_2^*$. Conversely, take curves $\gamma_1$ and $\gamma_2$ with parameter intervals $[\alpha_1,\beta_1]$ and $[\alpha_2,\beta_2]$. So long as $\gamma_1(\beta_1) = \gamma_2(\alpha_2)$ we can form the **join**, $\gamma$ say, of $\gamma_1$ and $\gamma_2$ (denoted $\gamma_1 \cup \gamma_2$). The recipe is

$$\gamma(t) := \begin{cases} \gamma_1(t) & \text{if } t \in [\alpha_1,\beta_1], \\ \gamma_2(t + \alpha_2 - \beta_1) & \text{if } t \in [\beta_1, \beta_1 + \beta_2 - \alpha_2]. \end{cases}$$

To avoid irritating technicalities later, the parameter intervals of $\gamma_1$ and $\gamma_2$ are here allowed to be arbitrary (whereas those of subintervals obtained by restriction automatically slot together). The penalty is a slightly complicated formula for join—essentially the parameter interval of $\gamma_2$ has to be translated. The joining process can be iterated: the join of $\gamma_1, \gamma_2, \ldots, \gamma_n$ can be defined provided the initial point of $\gamma_k$ coincides with the final point of $\gamma_{k-1}$ for $k = 1, \ldots, n-1$. A polygonal route as defined in 3.11 is the image of the join of line segments viewed as curves.

A curve $\gamma$ is said to be **smooth** if the function $\gamma$ has a continuous derivative on its parameter interval $[\alpha,\beta]$ (derivatives at $\alpha$ and $\beta$ being one-sided). A **path** is the join of finitely many smooth curves.

In the illustrative diagrams in Fig. 4.1 we perforce depict $\gamma^*$ (the image) rather than $\gamma$ (the function). Arrows indicate the direction in which $\gamma^*$ is traced.

It should be noted that, even when $\gamma$ is a path, $\gamma^*$ may be extremely complicated, to an extent that diagrams cannot adequately convey.

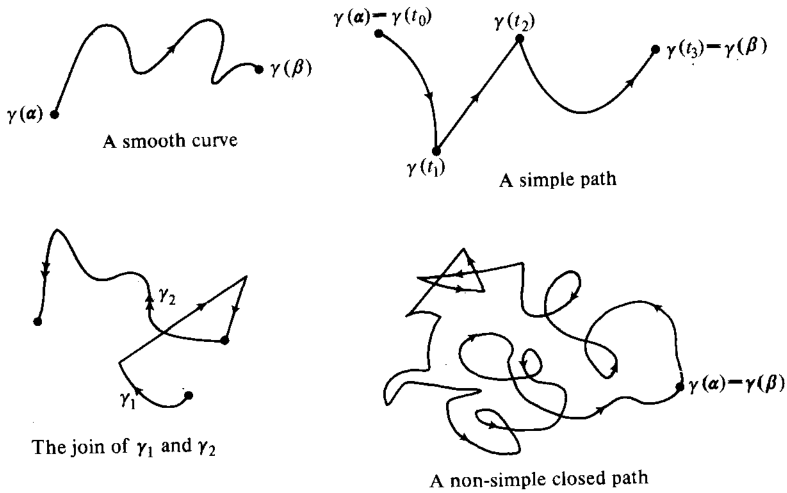



**Figure 4.1** Curves and paths

**4.3 Circline paths and contours.** Familiar figures such as circles and squares can be realized as images of paths. In particular:

- for any $u$ and $v$ in $\mathbb{C}$, the image of the path given by

$$\gamma(t) = (1-t)u + tv \quad (t \in [0,1])$$

  is the line segment $[u, v]$ traced from $u$ to $v$, and

- any circular arc traced clockwise (or anticlockwise) is the image of a path $\gamma$ (or $-\gamma$), where

$$\gamma(t) = a + re^{it} \quad (t \in [\theta_1, \theta_2]),$$

  with $a \in \mathbb{C}$, $r > 0$, and $0 \leqslant \theta_2 - \theta_1 \leqslant 2\pi$.

We define a **circline path** to be a path which is the join of finitely many paths of these types and a **contour** to be a simple closed circline path. The image of a contour consists of finitely many line segments and circular arcs and does not cross itself. We define a contour $\gamma$ to be **positively oriented** if, as $t$ increases,

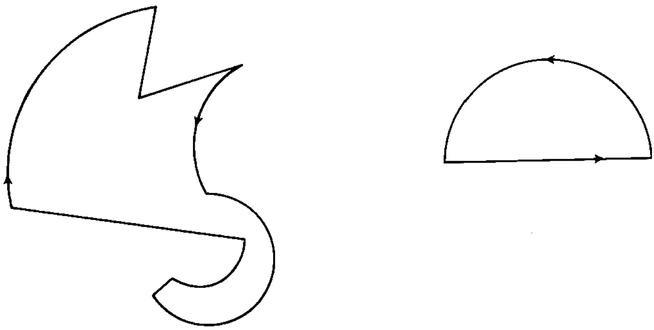


**Figure 4.2** Two contours

$\gamma(t)$ moves anticlockwise round any point inside it. [A more formal definition, in terms of index, is given in 12.8, on the advanced track.]

A geometric adjective (circular, triangular, ...) applied to a path or contour $\gamma$ will refer to the shape of its image $\gamma^*$. However we shall shorten 'triangular contour' to 'triangle', etc., where such an abuse of terminology will not cause confusion. We introduce the following notation: $\gamma(a;r)$ denotes the circle centre $a$ and radius $r$ given by

$$\gamma(a;r)(t) := a + r\mathrm{e}^{\mathrm{i}t} \quad (t \in [0, 2\pi]),$$

$\Gamma_r(t) := r\mathrm{e}^{\mathrm{i}t} \quad (t \in [0,\pi])$ defines a frequently used circular arc, and $[u, v]$ denotes a line segment regarded either as a path or as a subset of $\mathbb{C}$.

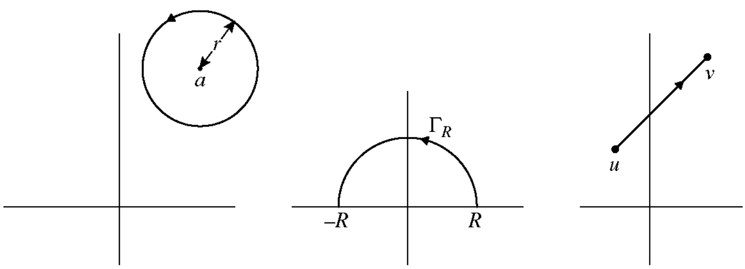



**Figure 4.3** Some particular circline paths

You should be aware that the term 'contour' is customarily used in a wider sense. As we have defined them, contours encompass all the paths regularly arising in applied complex analysis. They have the virtue that their images are geometrically much simpler than those of arbitrary closed paths. Even for contours, the geometric properties we require—the existence of an 'inside' and an 'outside', for example—though obvious in most specific cases, are tricky to prove in general. For convenience, proofs of all results of this kind are collected together in the next section.

## Properties of paths and contours

The results in this section are not needed until later. Since all are highly plausible, some readers may be content to take the conclusions on trust and to skip over the proofs. It is, however, instructive to see what is involved in proving the 'obvious'. In justifying geometric statements we have opted for an indication of strategy at the expense of detail.

Our first goal is the Covering theorem. The theorem asserts that, if $\gamma$ is a path lying in an open set $G$, then its image $\gamma^*$ can be covered by a finite chain of open sets contained in $G$, each overlapping the next. Our strategy is first to show that we can cover $\gamma^*$ with discs all of the same radius and then to show that only finitely many of these discs are needed. [The Covering theorem is closely related to the Heine–Borel theorem, applied to $\gamma^*$.]

**4.4 Lemma.** Let $\gamma$ be a path lying in an open set $G$. Then there exists a constant $m > 0$ such that $\mathrm{D}(z;m) \subseteq G$ for all $z \in \gamma^*$.

**Proof** Suppose, for a contradiction, that no such $m$ exists. Then for each $n$ there exist $z_n \in \gamma^*$ and $w_n \notin G$ such that $|z_n - w_n| < 1/n$. Use the Convergent subsequence theorem to choose a subsequence $\{z_{n_r}\}$ of $\{z_n\}$ which converges, to $z$ say. Then $w_{n_r} \to z$ too and so, since $\mathbb{C} \setminus G$ is closed, $z \notin G$. But we also have $z \in \gamma^*$, since $\gamma^*$ is a closed set. We therefore have the required contradiction. □

**4.5 Covering theorem.** Suppose that $G$ is an open set and that $\gamma$ is a path with parameter interval $[\alpha, \beta]$ such that $\gamma^* \subseteq G$. Then there exist a constant $m > 0$ and open discs $D_0, D_1, \ldots, D_N$ such that

(i) for $k = 0, 1, \ldots, N$, $D_k = \mathrm{D}(\gamma(t_k); m)$, where $\alpha = t_0 < t_1 < \cdots < t_N = \beta$;
(ii) for $k = 0, \ldots, N-1$, $D_k \cap D_{k+1} \neq \varnothing$;
(iii) for $k = 0, \ldots, N-1$, $\gamma([t_k, t_{k+1}]) \subseteq D_k$;
(iv) $\gamma^* \subseteq \bigcup_{k=0}^{N} D_k \subseteq G$.

(The disc $D_N$ is not needed for the covering. It is put in for later notational convenience. When $\gamma$ is closed, $D_0$ and $D_N$ coincide.)

**Proof** Choose $m$ as in Lemma 4.4, so that $D(z;m) \subseteq G$ for every $z \in \gamma^*$. It remains to show that $\gamma^*$ can be covered by a *finite* chain of such discs, each overlapping the next, as in Fig. 4.4. If $\gamma^*$ is made up of finitely many line segments and circular arcs (in particular, if $\gamma$ is a contour) then this is clear from elementary geometry.

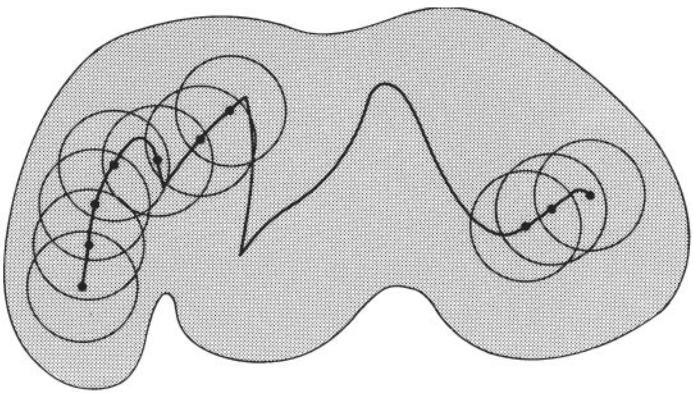


**Figure 4.4** The Covering theorem

For a general path we proceed as follows. Suppose first that $\gamma$ is smooth. Then the real Mean value theorem implies that, for any $s$ and $t$ in $[\alpha, \beta]$,

$$(\operatorname{Re}\gamma)(s) - (\operatorname{Re}\gamma)(t) = (s-t)(\operatorname{Re}\gamma)'(c)$$

for some $c$ between $\alpha$ and $\beta$, and similarly for $\operatorname{Im}\gamma$. The continuous functions $(\operatorname{Re}\gamma)'$ and $(\operatorname{Im}\gamma)'$ are bounded on $[\alpha, \beta]$, by 3.24. Therefore there exists $\delta > 0$ such that

$$|\gamma(s) - \gamma(t)| < m \quad \text{whenever } |s-t| < \delta.$$

This conclusion [uniform continuity of $\gamma$ to those in the know] persists for an arbitrary path, since we can apply the above argument to its constituent smooth curves.

We can now select points $\alpha = t_0 < t_1 < \cdots < t_N = \beta$ satisfying $|t_{k+1} - t_k| < \delta$ for $k = 0, 1, \ldots, N-1$. If we choose $D_k$ to be $D(\gamma(t_k); m)$ for $k = 0, 1, \ldots, N$ then conditions (i)–(iv) of the theorem are met. □

**Remark** By refining the argument above it can be shown that a little more is true than we claimed: for some $\eta > 0$, the open strip $S := \{z : \exists w \in \gamma^* \text{ with } |z - w| < \eta\}$ is such that $\gamma^* \subseteq S \subseteq G$.)

The famous Jordan curve theorem asserts that a simple closed path has an 'inside' and an 'outside'. In its general form, it is a very deep result. We restrict attention to contours (as defined in 4.3).

**4.6 Jordan curve theorem (for a contour).** Let $\gamma$ be a contour. Then the complement of $\gamma^*$ is of the form $\mathbf{I}(\gamma) \cup \mathbf{O}(\gamma)$, where $\mathbf{I}(\gamma)$ and $\mathbf{O}(\gamma)$ are disjoint connected open sets, $\mathbf{I}(\gamma)$ (the **inside** of $\gamma$) is bounded and $\mathbf{O}(\gamma)$ (the **outside** of $\gamma$) is unbounded.

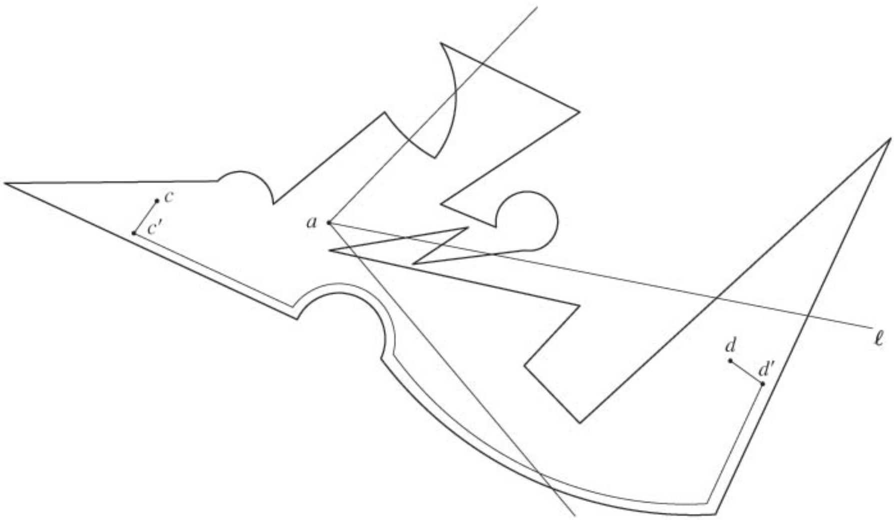



**Figure 4.5** Connectedness of $\mathbf{I}(\gamma)$

**Outline proof.** (For further details, consult Kosnioswki [**17**], pp. 102–103.) For any fixed $a \notin \gamma^*$, consider a ray $\ell$ with endpoint $a$. Let $N(a, \ell)$ be the number of times $\ell$ cuts $\gamma^*$ (this is well defined except for, at worst, finitely many positions of $\ell$ involving tangency or 'corner points'; we leave $N(a, \ell)$ undefined in these degenerate cases). The crucial point to note is that whether $N(a, \ell)$ is odd or even depends only on $a$ and not on the direction of $\ell$; see Fig. 4.5. Let $\mathbf{I}(\gamma)$ ($\mathbf{O}(\gamma)$) consist of those points $a \notin \gamma^*$ for which $N(a, \ell)$ is odd (even).

That $\mathbf{I}(\gamma)$ and $\mathbf{O}(\gamma)$ are open follows from the observation that for $z \notin \gamma^*$ there is an open disc $\mathrm{D}(z; r)$ disjoint from $\gamma^*$ which lies wholly in $\mathbf{I}(\gamma)$ or wholly in $\mathbf{O}(\gamma)$ (given $w \in \mathrm{D}(z; r)$, consider $[w, z]$ extended to a ray with endpoint $z$). To prove connectedness of, say, $\mathbf{I}(\gamma)$ it is sufficient to show that any two points $c$

and $d$ in $\mathbf{I}(\gamma)$ can be joined by a path in $\mathbf{I}(\gamma)$ made up of line segments and circular arcs (see 3.14). Figure 4.5 shows how this can be done: we join $c$ and $d$ to points $c'$ and $d'$ in $\mathbf{I}(\gamma)$ close to $\gamma^*$ and, following $\gamma^*$ at a fixed small positive distance from it and staying within $\mathbf{I}(\gamma)$, connect $c'$ to $d'$. □

**4.7 Boundaries.** Our intuition on what constitutes the boundary of a set $S \subseteq \mathbb{C}$ is only adequate if $S$ is bounded by some familiar geometric figure. But subsets of the plane can be geometrically very complicated indeed, so we need a formal, topological, definition. The **boundary of a set** $S$ is $\partial S := \overline{S} \cap \overline{\mathbb{C} \setminus S}$. When $S$ is open, $\mathbb{C} \setminus S$ is closed, and so equal to its own closure. In this case, $\partial S = \overline{S} \setminus S$.

For a contour $\gamma$, both $\gamma^* \cup \mathbf{I}(\gamma)$ and $\gamma^* \cup \mathbf{O}(\gamma)$ are closed and the set $\gamma^*$ is the boundary both of $\mathbf{I}(\gamma)$ and of $\mathbf{O}(\gamma)$.

Integration of functions along paths is the subject of Chapter 10. Our final result about paths will allow us to break up an integral round a polygonal contour into a sum of integrals round triangles, a crucial step in the proof of Cauchy's theorem I (11.6).

**4.8 Triangulation of a polygon.** Let $\gamma$ be a polygonal contour in $\mathbb{C}$ and let $z_1, z_2, \ldots, z_n$ $(n > 3)$ be the vertices of $\gamma^*$. Then it is possible to insert $n - 3$ line segments $[z_j, z_k]$ so as to subdivide $\mathbf{I}(\gamma)$ into $n - 2$ triangles. Each of the inserted segments, excluding its endpoints, lies in $\mathbf{I}(\gamma)$.

**Outline proof.** (For further details consult Hille **[14]**, p. 286.) If $\mathbf{I}(\gamma)$ is convex, then the segments $[z_1, z_k]$, $k = 3, \ldots, n-1$, triangulate it. Otherwise the interior angle at some vertex, say $z_1$, is greater than $\pi$. Consider a ray $\ell$ emanating from $z_1$ such that $\mathrm{D}(z_1; r) \cap \mathbf{I}(\gamma) \neq \varnothing$ for all $r$ sufficiently small (so $\ell$ points into $\mathbf{I}(\gamma)$). Moving along such a ray from $z_1$, there is a first point of intersection $w_\ell$ $(\neq z_1)$ of $\ell$ with $\gamma^*$. For at least one choice of $\ell$, the point $w_\ell$ is a vertex of the polygon. Let $z_k$ be such a vertex. The segment $[z_1, z_k]$ can then be used to create two new polygonal contours, each of whose images in $\mathbb{C}$ has fewer than $n$ vertices. The argument is repeated until only triangles remain. The process is illustrated in Fig. 4.6. □

## Exercises

**Exercises from the text.** Prove that every point of $\gamma^*$ belongs to $\overline{\mathbf{I}(\gamma)}$ and to $\overline{\mathbf{O}(\gamma)}$, for any contour $\gamma$ (see 4.6).

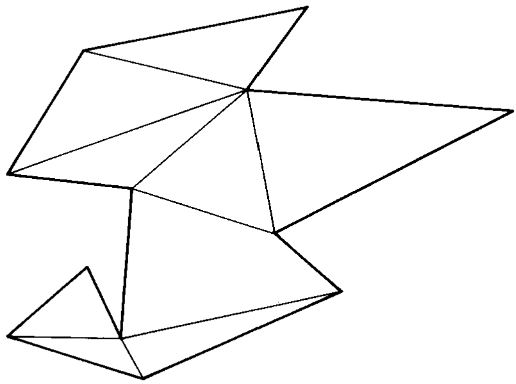


**Figure 4.6** Triangulation of a polygon

4.1 Describe the image $\gamma^*$ of the curve $\gamma$ in the following cases, indicating how the image is traced.

(i) $\gamma(t) = 1 + \mathrm{i}e^{\mathrm{i}t} \quad (t \in [0, \pi])$.

(ii) $\gamma(t) = e^{\mathrm{i}t} \quad (t \in [-\pi, 2\pi])$.

(iii) $\gamma$ is the join of $[-1, 1]$, $[1, 1+\mathrm{i}]$, $[1+\mathrm{i}, -1-\mathrm{i}]$.

(iv) $\gamma$ is the join of $\gamma_1$, $\gamma_2$, $\gamma_3$, where $\gamma_1$ is $[1-\mathrm{i}, 0]$, $\gamma_2$ is $[0, 1+\mathrm{i}]$, and $\gamma_3$ is defined by $\gamma_3(t) = \sqrt{2}e^{\mathrm{i}(t+\frac{1}{4}\pi)}$ $(t \in [0, 3\pi/2])$.

(v) $\gamma$ has parameter interval $[0, 2\pi]$ and is given by $\gamma(t) = e^{\mathrm{i}t}$ for $0 \leqslant t \leqslant \pi$ and by $\gamma(t) = e^{-\mathrm{i}t}$ for $\pi \leqslant t \leqslant 2\pi$.

(vi) $\gamma(t) = e^{\mathrm{i}t}\cos t$ $(t \in [0, 2\pi])$.

In which cases is $\gamma$ (a) closed, (b) simple, (c) smooth, (d) a path?

4.2 Define parametrically a path $\gamma$ for which $\gamma^*$ is

(i) the square with vertices at $\pm 1 \pm \mathrm{i}$;

(ii) the closed semicircle in the right half-plane with $[-R\mathrm{i}, R\mathrm{i}]$ as diameter;

(iii) the pair of circles $|z-1| = 1$ and $|z+1| = 1$, the first traced clockwise and the second anticlockwise.

# 5 Holomorphic functions

Complex analysis may be summarized as the study of holomorphic functions. Holomorphic means—almost—the same as differentiable, but there is a critical distinction between the two concepts. This comes from the role played by open sets.

## Differentiation and the Cauchy–Riemann equations

**5.1 Differentiation.** A complex-valued function $f$ defined on an *open* subset $G$ of $\mathbb{C}$ is **differentiable** at $z \in G$ if

$$\lim_{h \to 0} \frac{f(z+h) - f(z)}{z}$$

exists. When the limit does exist it is denoted by $f'(z)$.

Note carefully the role of the open set here. Since $G$ is open, we know that, given $z \in G$, there exists $r > 0$ such that $\mathrm{D}(z;r) \subseteq G$. That is, $z + h \in G$ whenever $|h| < r$. So, in the computation of the limit defining $f'(z)$, the point $z+h$ is free to approach $z$ from any direction as $h \to 0$. For $f$ to be differentiable at $z$ it is necessary that the quotient $(f(z+h)-f(z))/h$ tend to a limiting value independent of the manner in which $h \to 0$. Turning this around, we see in particular that $f$ cannot be differentiable at $z$ if the quotient $(f(z+h)-f(z))/h$ has different limiting values when $h$ approaches 0 from different directions. A similar situation can arise in real analysis: a real-valued function defined on an open interval of $\mathbb{R}$ is not differentiable at $x$ if its left-hand and right-hand derivatives at $x$ exist but have different values.

**5.2 Example (a non-differentiable function).** Let $f(z) = \mathrm{Re}\, z$ in $\mathbb{C}$. We show that $f$ is not differentiable at any point $z \in \mathbb{C}$:

$$\begin{aligned} \frac{f(z+h) - f(z)}{h} &= \frac{\mathrm{Re}\,(z+h) - \mathrm{Re}\, z}{h} \\ &= \frac{\mathrm{Re}\, h}{h} \to \begin{cases} 1 & \text{as } h \to 0 \text{ with } h \text{ real,} \\ 0 & \text{as } h \to 0 \text{ with } h \text{ purely imaginary.} \end{cases} \end{aligned}$$

Thus $f'(z)$ does not exist.

The idea of restricting $h$ to be real or purely imaginary can be exploited quite generally, and yields a necessary (but not sufficient) condition for differentiability.

**5.3 Theorem (the Cauchy–Riemann equations).** Let the complex-valued function $f$ be defined on an open set $G$ and be differentiable at $z = x + \mathrm{i}y \in G$. Let $f(z) = u(x, y) + \mathrm{i}v(x, y)$ (as in 1.10). Then $u$ and $v$ have first-order partial derivatives at $(x, y)$ (denoted $u_x, u_y, v_x, v_y$) and these satisfy the **Cauchy–Riemann equations**

$$u_x = v_y, \quad u_y = -v_x.$$

**Proof** From the definition in 5.1,

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$$

exists. Hence, restricting $h$ to be, respectively, real and purely imaginary, we have

$$f'(z) = \lim_{\substack{h \to 0 \\ h \in \mathbb{R}}} \frac{u(x+h, y) - u(x, y)}{h} + \mathrm{i}\frac{v(x+h, y) - v(x, y)}{h} = u_x + \mathrm{i}v_x$$

and

$$f'(z) = \lim_{\substack{h \to 0 \\ h = \mathrm{i}k \\ k \in \mathbb{R}}} \frac{u(x, y+k) - u(x, y)}{\mathrm{i}k} + \frac{v(x, y+k) - v(x, y)}{k} = \frac{1}{\mathrm{i}}u_y + v_y.$$

(The existence of the partial derivatives follows from 3.18(2).) Equating the two expressions for $f'(z)$ gives

$$u_x + \mathrm{i}v_x = -\mathrm{i}u_y + v_y.$$

Equating real and imaginary parts we obtain

$$u_x = v_y, \quad u_y = -v_x. \quad \square$$

**5.4 Examples (Cauchy–Riemann equations).**

- (Example 5.2 revisited) Consider $f(z) = \operatorname{Re} z$ on $G = \mathbb{C}$. Here we have $u(x, y) = x$ and $v(x, y) = 0$ so $u_x = 1 \neq 0 = v_y$. By (the contrapositive of) the Cauchy–Riemann equations, $f$ is not differentiable anywhere.

- Let $f(z) = |z|$ on $G = \mathbb{C}$. Here

$$u(x,y) = (x^2+y^2)^{1/2} \quad \text{and} \quad v(x,y) = 0.$$

Then $v_x = v_y = 0$ and, for $(x,y) \neq (0,0)$,

$$u_x = x(x^2+y^2)^{-1/2} \quad \text{and} \quad u_y = y(x^2+y^2)^{-1/2}.$$

The Cauchy–Riemann equations fail to hold, and so $f$ fails to be differentiable, at any point $z \neq 0$. The point 0 requires separate attention. From first principles,

$$\frac{f(h)-f(0)}{h} = \frac{|h|}{h} \to \begin{cases} 1 & \text{as } h \to 0 \text{ with } h \text{ real and positive,} \\ -1 & \text{as } h \to 0 \text{ with } h \text{ real and negative.} \end{cases}$$

Hence $f'(0)$ does not exist.

**5.5 The limitations of the Cauchy–Riemann equations.** The contrapositive of Theorem 5.3 is useful for proving *non*-differentiability. Conversely, the Cauchy–Riemann equations are *not* on their own sufficient to guarantee differentiability. Here is an easy but artificial example to show this. Let

$$f(z) = f(x,y) = \begin{cases} 1 & \text{if neither } x \text{ nor } y \text{ is zero,} \\ 0 & \text{otherwise,} \end{cases}$$

that is, $f$ takes the value 1 except on the $x$- and $y$-axes, and is zero there. At 0 we have $u_x = u_y = v_x = v_y = 0$, so the Cauchy–Riemann equations hold. However $\lim_{h\to 0}(f(h)-f(0))/h$ fails to exist, as we see, for example, by letting $h$ approach 0 along the ray $\arg z = \pi/4$. Other examples can be found in Exercise 5.5.

It turns out that, provided we impose continuity conditions on the partial derivatives, we do obtain a converse to Theorem 5.3. We record this technical result here for completeness, but we do *not* recommend its use as a practical means of testing for differentiability in open sets. Soon we shall have much better methods for establishing this. We do use the lemma in Chapter 23, where we study the relationship between harmonic functions (which are smooth solutions of Laplace's equation in two dimensions, $u_{xx} + u_{yy} = 0$) and holomorphic functions.

**5.6 Technical lemma (a partial converse to Theorem 5.3).** Let $f(z) = u(x,y) + \mathrm{i}v(x,y)$ for $z = x + \mathrm{i}y \in G$, where $G \subseteq \mathbb{C}$ is open. Assume that $u$ and $v$ have continuous first-order partial derivatives in $G$ and that they satisfy the Cauchy–Riemann equations at $z$. Then $f'(z)$ exists.

**Proof** Let $z \in G$ and choose $r$ such that $\mathrm{D}(z;r) \subseteq G$. Take $h = p + \mathrm{i}q$ with $|h| < r$. Then

$$\begin{aligned}
&\frac{f(z+h) - f(z)}{h} \\
&\quad = \frac{p}{h}\left(\frac{u(x+p, y+q) - u(x, y+q)}{p} + \mathrm{i}\frac{v(x+p, y+q) - v(x, y+q)}{p}\right) \\
&\qquad\qquad + \frac{q}{h}\left(\frac{u(x, y+q) - u(x,y)}{q} + \mathrm{i}\frac{v(x, y+q) - v(x,q)}{q}\right) \\
&\quad = \frac{p}{h}\left(\frac{\partial u}{\partial x}(x + \alpha p, y+q) + \mathrm{i}\frac{\partial v}{\partial x}(x + \beta p, y+q)\right) \\
&\qquad\qquad + \frac{q}{h}\left(\frac{\partial u}{\partial y}(x, y + \gamma q) + \mathrm{i}\frac{\partial v}{\partial y}(x, y + \delta q)\right),
\end{aligned}$$

where each of $\alpha, \beta, \gamma, \delta$ lies in $(0,1)$. Here we have applied the real Mean value theorem to the four functions $x \mapsto u(x, y+p)$, $x \mapsto v(x, y+p)$, $y \mapsto u(x,y)$, $y \mapsto v(x,y)$. Using the continuity of the partial derivatives and the Cauchy–Riemann equations we see that $f'(z)$ exists. □

## Holomorphic functions

In the preceding section we linked differentiation with respect to $z$ to partial differentiation with respect to the real variables $x$ and $y$. We now cut loose from real analysis and work directly with a complex variable.

**5.7 Definition (holomorphic function).** We have already seen the merits of working in an open set. In defining holomorphy, open sets play an integral part.

(1) A complex-valued function $f$ which is differentiable at every point of an open set $G$ is said to be **holomorphic in** $G$. We reiterate that this means that $\lim_{h \to 0}(f(z+h) - f(z))/h$ exists (independently of the manner in which $h$ approaches 0) for each $z \in G$. The set of functions holomorphic in $G$ is denoted $\mathrm{H}(G)$.

(2) A complex-valued function $f$ is said to be **holomorphic at a point** $a \in \mathbb{C}$ if there exists $r > 0$ such that $f$ is defined and holomorphic in $\mathrm{D}(a;r)$.

We stress that being holomorphic at a point $a$ is a stronger condition than simply being differentiable at $a$.

**5.8 Holomorphic functions: elementary properties.** We have so far refrained from giving examples of differentiable functions because, except in the simplest cases, it is laborious to check differentiability direct from the definition. (As in real analysis, proving the existence of a derivative from first principles is mainly an exercise on limits.) We have also advised against using Technical lemma 5.6 to prove differentiability. Instead, we build up a catalogue of holomorphic functions by forming products, composites, etc.

Let $G$ be a fixed open set. The following properties are proved by checking the appropriate differentiability conditions at each point of $G$. We omit the details as the proofs are formally identical to their real counterparts.

(1) Let $f$ and $g$ be holomorphic in $G$ and let $\lambda \in \mathbb{C}$. Then $\lambda f$, $f+g$, and $fg$ (all defined pointwise in the usual way) are holomorphic in $G$ and the usual differentiation rules apply: for all $z \in G$,

$$\begin{aligned}(\lambda f)'(z) &= \lambda f'(z),\\ (f+g)'(z) &= f'(z)+g'(z), \text{ and}\\ (fg)'(z) &= f'(z)g(z)+f(z)g'(z).\end{aligned}$$

(2) **Chain rule** Let $f$ be holomorphic in $G$ and let $g$ be holomorphic in an open set containing $f(G)$. Then the composite function $g \circ f$, given by $(g \circ f)(z) = g(f(z))$, is holomorphic in $G$ and, for all $z \in G$,

$$(g \circ f)'(z) = g'(f(z))f'(z).$$

(3) Let $f$ be holomorphic in $G$ and suppose that $f(z) \neq 0$ for all $z \in G$. Then $1/f$ is holomorphic in $G$ and, for any $z \in G$,

$$(1/f)'(z) = f'(z)/(f(z))^2.$$

**5.9 Holomorphic functions: preliminary examples.** We can now easily construct examples of holomorphic functions. The function $f$ defined by $f(z) = z$ is certainly differentiable everywhere, as is any constant function. By 5.8(1), any **polynomial**

$$p(z) = \sum_{n=0}^{N} c_n z^n \quad (c_n \in \mathbb{C},\ N \text{ an integer } \geqslant 0)$$

is holomorphic in $\mathbb{C}$. We emphasize that a polynomial is a *finite* sum of terms of the form $c_n z^n$. The corresponding infinite sums—power series—are centrally important but raise issues of convergence. They do not come on stage until the next chapter.

By 5.8(1) & (3), a **rational function** $p(z)/q(z)$ ($p(z)$ and $q(z)$ polynomials) is holomorphic in any open set in which $q(z)$ is never zero. For example, $(1+z^2)^{-2}$ is holomorphic in the open set $\mathbb{C} \setminus \{\pm \mathrm{i}\}$ (and so holomorphic at every point except i and $-$i).

**5.10 Behaviour of functions at $\infty$.** We have already briefly considered functions on the extended plane $\widetilde{\mathbb{C}}$ in connection with Möbius transformations. We can use the inversion map $z \mapsto 1/z$ to analyse what happens at or near $\infty$. Consider a function $f$ defined on some set $\{ z \in \mathbb{C} : |z| > r \}$ but not necessarily at the point $\infty$. Define $\widetilde{f}$ by

$$\widetilde{f}(z) = f(1/z) \quad (z \in \mathrm{D}'(0; 1/r))$$

and let $\widetilde{f}(0) = f(\infty)$ if $f(\infty)$ is defined. We then transfer notions relating to $\widetilde{f}$ at 0 to obtain corresponding notions for $f$ at $\infty$: limiting value, continuous, holomorphic, and so on.

For example, consider $f(z) = z^3$. Then we have $\widetilde{f}(w) = w^{-3}$, and this is not holomorphic at $w = 0$. Now let $f(z) = (1 + z^2)^{-1}$ for $|z| > 1$ and let $f(\infty) = 0$. Then $\widetilde{f}(w) = w^2(w^2 + 1)^{-1}$ for $|w| < 1$. Therefore $f$ is holomorphic at $\infty$.

The extended plane is the right setting for studying Möbius transformations. Consider $f : \widetilde{\mathbb{C}} \to \widetilde{\mathbb{C}}$ given by $f(z) = (az + b)/(cz + d)$ $(ad - bc \neq 0)$. We may regard $f$ as a rational function; as such, it is holomorphic in $\mathbb{C}$ except at $-d/c$ if $c \neq 0$ and everywhere in $\mathbb{C}$ if $c = 0$.

Now assume $c \neq 0$. The arithmetic rules in 2.8 imply that $f(\infty) = a/c$ (divide by $z$ top and bottom before putting $z = \infty$) and that $\widetilde{f}$ is given by $\widetilde{f}(w) = (a + bw)/(c + dw)$. Certainly $\widetilde{f}$ is holomorphic at $w = 0$, so $f$ is holomorphic at $\infty$.

The remaining results in this chapter are ones we shall use frequently in an ancillary role.

**5.11 Holomorphy implies continuity.** We shall establish the technically useful fact that, if $f$ is holomorphic in $G$, then $f$ is continuous on $G$. Combining this with Theorem 3.24, we can then assert that if, also, $S$ is a compact subset of $G$ (that is, $S$ is closed and bounded) then $f$ is bounded on $S$.

To prove our claim, suppose that $f$ is differentiable at a point $z$ of an open set $G$ in which $f$ is defined. For $h$ such that $z+h \in G$,

$$f(z+h) = f(z) + hf'(z) + h\varepsilon(h) \quad \text{where } \varepsilon(h) \to 0 \text{ as } h \to 0.$$

To see this, we write, for $h \neq 0$,

$$\varepsilon(h) := \frac{f(z+h) - f(z)}{h} - f'(z).$$

It follows that $f(z+h) \to f(z)$ as $h \to 0$, as required.

The Cauchy–Riemann equations have useful theoretical consequences. The proof of the next proposition is unaesthetic, but instructive. Like that in 5.6, it provides a bridge between complexified real analysis and complex analysis proper.

**5.12 Proposition (constancy in a region).** Suppose that $f$ is holomorphic in a region $G$. Then any of the following conditions forces $f$ to be constant in $G$:
(1) $f'(z) = 0$ for all $z \in G$;
(2) $|f|$ constant in $G$;
(3) $f(z)$ is real for all $z \in G$.

**Proof** We first assume that $G$ is the unit disc $\mathrm{D}(0;1)$. We adopt the notation of Theorem 5.3. The proof of this theorem shows that, for $z = x + \mathrm{i}y \in \mathrm{D}(0;1)$,

$$f'(z) = u_x + \mathrm{i}v_x = -\mathrm{i}u_y + v_y.$$

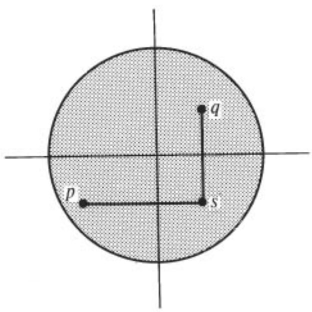



**Figure 5.1** Proof of Proposition 5.12 for $\mathrm{D}(0;1)$

Suppose that $f'$ is identically zero. Then $u_x = v_x = u_y = v_y$ throughout $D(0;1)$. Fix points $p = a + ib$ and $q = c + id$ in $D(0;1)$. We shall prove (1) by showing that $f(p) = f(q)$. At least one of $s = c + ib$ and $t = a + id$ lies in $D(0;1)$; suppose without loss of generality that $s$ does. Each of $x \mapsto u(x,b)$ and $y \mapsto u(c,y)$ is a real-valued function of a real variable with zero derivative, and so is constant, by the Mean value theorem. Hence

$$u(a,b) = u(c,b) \quad \text{and} \quad u(c,b) = u(c,d)$$

and likewise

$$v(a,b) = v(c,b) \quad \text{and} \quad v(c,b) = v(c,d).$$

We conclude that $f(p) = f(s) = f(q)$.

Now consider (2). Suppose that $|f(z)| = c$ for $|z| < 1$, where $c$ is a constant. We then have $u^2 + v^2 = c^2$. Therefore

$$uu_x + vv_x = 0, \quad uu_y + vv_y = 0.$$

Hence, by the Cauchy–Riemann equations,

$$uu_x - vu_y = 0, \quad uu_y + vu_x = 0.$$

Elimination of $u_y$ gives $0 = (u^2 + v^2)u_x = c^2 u_x$. If $c = 0$ then we have $f \equiv 0$, trivially. Otherwise, $u_x = 0$ everywhere in $D(0;1)$. Similarly, $u_y$, $v_x$, and $v_y$ are zero. We deduce, from above, that $f$ is constant.

The proof of (3) is similar. If $f$ is real-valued, then $v = 0$, so that $v_x = v_y = 0$. By the Cauchy–Riemann equations, $u_x = u_y = 0$ too. Hence, as before, $f$ must be constant.

To extend the proof of (1) (and hence of (2) and (3) too) to the case that $G$ is an arbitrary region, we appeal to 3.14. Any two points in $G$ can be joined by a polygonal route consisting of horizontal and vertical line segments. The proof is now a mild complication of that given earlier. □

**5.13 Functions with zero derivative: postscript.** Consider the open set $G = D(-2;1) \cup D(2;1)$ and define $f$ on $G$ by

$$f(z) = \begin{cases} 1 & \text{if } z \in D(-2;1), \\ -1 & \text{if } z \in D(2;1). \end{cases}$$

Then, working within the two open discs separately, we see that $f' = 0$ in $G$. More generally, let $G$ be a non-empty open set which is not connected. Then $G$

can be partitioned into disjoint open sets $G_1$ and $G_2$, and any function defined to take different constant values on $G_1$ and $G_2$ has zero derivative but is non-constant.

**5.14 Beware! Non-holomorphic functions at large.** Failure of the Cauchy–Riemann equations signals non-differentiability. It happens, in particular, for any non-constant *real-valued* function in an open disc or, more generally, in a region (see 5.12). As a consequence, various functions derived from a non-constant function $f = u + iv$ which is holomorphic in a region $G$ cannot themselves be holomorphic. For example, none of

$$|f|, \quad u = \operatorname{Re} f, \quad v = \operatorname{Im} f$$

is holomorphic anywhere. Contrast this with the situation as regards continuity: when $f$ is continuous, $|f|$, $\operatorname{Re} f$, and $\operatorname{Im} f$ are also continuous.

We also note that any function which is differentiable just at a single point, or just on some set of isolated points, cannot be holomorphic anywhere. The reason is that, by definition, holomorphic at a point $a$ means differentiable at every point of some disc $\mathrm{D}(a; r)$ $(r > 0)$.

## Exercises

5.1 Verify that (i) $\operatorname{Im} z$ and (ii) $\bar{z}$ do not satisfy the Cauchy–Riemann equations at any point $z = x + iy$ in $\mathbb{C}$ (so neither function is differentiable anywhere).

5.2 Verify directly that the functions given in Exercise 1.15 satisfy the Cauchy–Riemann equations.

5.3 The derivation of the Cauchy–Riemann equations in Theorem 5.3 shows that, if $f = u + iv$ is differentiable at $z$, then $f'(z) = u_x + iv_x = -iu_y + v_y$. Verify that

$$f'(z) = u_x - iu_y \quad \text{and} \quad f'(z) = v_y + iv_x.$$

(So $f'$ is determined by either of $u$ or $v$ alone.)

5.4 Which of the following functions is differentiable at $z = 0$? Give a proof or refutation as appropriate.

$$\text{(i) } |z|^2, \quad \text{(ii) } \operatorname{Re} z + \operatorname{Im} z, \quad \text{(iii) } (\operatorname{Re} z)(\operatorname{Im} z).$$

5.5 (a) Prove that $f$ defined by

$$f(z) = \begin{cases} z^5/|z|^4 & \text{if } z \neq 0, \\ 0 & \text{if } z = 0 \end{cases}$$

satisfies the Cauchy–Riemann equations at $z = 0$ but is not differentiable there.

(b) Prove that $f$ defined by

$$f(z) = \sqrt{|(\operatorname{Re} z)(\operatorname{Im} z)|}$$

satisfies the Cauchy–Riemann equations at $z = 0$ but is not differentiable there.

5.6 (This exercise shows that the Mean value theorem from real analysis does not have a direct complex analogue.) Let $f(z) = z^3$. Prove that there exists no point $c$ on the line segment $[1, \mathrm{i}]$ such that

$$\frac{f(\mathrm{i}) - f(1)}{\mathrm{i} - 1} = f'(c).$$

5.7 (a) At which points $z = x + \mathrm{i}y$ are the following functions holomorphic?

(i) $z^8 + 7z^5 - \pi z^2 + 1$, (ii) $(z(z-1)(z-2))^{-2}$, (iii) $(z^5 - 1)^{-1}$.

(b) Prove that the following are not holomorphic at any point:

(i) $1/|z|$, (ii) $z|z|$.

5.8 Prove that $z/(1 + |z|)$ is not holomorphic anywhere (cf. Exercise 3.13). (Hint: argue by contradiction and exploit the fact that $|z|$ is differentiable nowhere rather than showing that the Cauchy–Riemann equations do not hold.)

5.9 Give examples of

(i) a function holomorphic except at $\pm 1$;

(ii) a function $f$ holomorphic in $\mathbb{C}$ for which $1/f$ fails to be holomorphic at precisely six points of $\mathbb{C}$;

(iii) a function $f = u + \mathrm{i}v$ for which neither $u$ nor $v$ is constant and which is holomorphic nowhere.

5.10 Suppose that $f$ is holomorphic in a region $G$. Prove that $f$ is constant if $\operatorname{Re} f$ is constant. .

5.11 Let $f$ be holomorphic in $D(0;1)$.

(i) Define $g$ by $g(z) = \overline{f(\overline{z})}$. Prove that $g$ is holomorphic in $D(0;1)$. (Hint: consider $\lim_{h\to 0}(g(z+h)-g(z))/h$.)

(ii) Define $k$ by $k(z) = \overline{f(z)}$. Prove that $k$ is differentiable at $a \in D(0;1)$ if and only if $f'(a) = 0$. Deduce that $k$ is holomorphic in $D(0;1)$ if and only if $f$ is constant.

5.12 Let $f$ be a complex-valued function which is differentiable at each point of an open set $G$. Define, for $z = (x,y) \in G$,

$$\frac{\partial f}{\partial \overline{z}} := \frac{1}{2}\left(\frac{\partial f}{\partial x} + \mathrm{i}\frac{\partial f}{\partial y}\right) \quad \text{and} \quad \frac{\partial f}{\partial z} := \frac{1}{2}\left(\frac{\partial f}{\partial x} - \mathrm{i}\frac{\partial f}{\partial y}\right),$$

where, on the right-hand side, $f$ is regarded as a function of $(x,y)$. Verify that the partial derivatives of $f$ with respect to $x$ and $y$ exist (so that the definitions above are valid ones) and show that

$$\frac{\partial f}{\partial \overline{z}} = 0 \quad \text{and} \quad \frac{\partial f}{\partial z} = f'.$$

Prove, conversely, that a differentiable function $f$ which satisfies $\partial f/\partial \overline{z} = 0$ in $G$ is holomorphic in $G$.

# 6 Complex series and power series

We begin this chapter with some introductory remarks to motivate the study of complex power series. As we shall see in Chapter 14, such series turn out to be fundamental in the theory of holomorphic functions.

In elementary calculus, use is often made of what are known as **Maclaurin expansions**: series expansions of the form

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \ldots.$$

Examples are

$$\begin{aligned} e^x &= 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \ldots, \\ \sin x &= x - \frac{1}{3!}x^3 + \cdots + (-1)^n \frac{1}{(2n+1)!}x^{2n+1} + \ldots, \\ \cosh x &= 1 + \frac{1}{2!}x^2 + \cdots + \frac{1}{(2n)!}x^{2n} + \ldots, \end{aligned}$$

for real $x$. We should like to have complex analogues of these. In the complex case, geometric ways of defining trigonometric functions are no longer available. We shall therefore wish to use series expansions to *define* functions such as sine and cosine.

There are two issues we must address if we are to use expansions

$$f(z) = f(0) + f'(0)z + \cdots + \frac{f^{(n)}(0)}{n!}z^n + \ldots$$

or, more generally,

$$f(z) = f(a) + f'(a)(z-a) + \cdots + \frac{f^{(n)}(a)}{n!}(z-a)^n + \ldots,$$

in the complex case. One is the existence of the derivatives $f^{(n)}$ and the other is convergence of the series. In real analysis, Taylor's theorem provides expansions of the form

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(N)}(a)}{N!}(x-a)^N + E_N,$$

$$\text{where } E_N = \frac{f^{(N+1)}(a+\theta_x x)}{(N+1)!}(x-a)^{N+1} \quad (\text{for some } |\theta_x| < 1),$$

for suitably well-behaved functions $f\colon \mathbb{R} \to \mathbb{R}$. Note that we do not have an infinite series here. It is also not essential for $f$ to have derivatives of all orders if our intention is to treat the Taylor expansion as an approximation $\sum_{n=0}^{N} f^{(n)}(a)(x-a)^n/n!$, with error $E_N$. On the other hand, it would certainly be preferable to have an infinite series expansion. This could be truncated after any desired number of terms to provide an estimate if required.

The situation in complex analysis is *much* better than that in real analysis. It turns out that a function which is holomorphic in an open set $G$ is such that $f^{(n)}(z)$ exists for $z \in G$ *for all* $n$: no worries about higher-order derivatives not necessarily existing! In addition, we shall see that there is a close and very satisfactory relationship between holomorphic functions and functions represented in open discs by power series (see Theorems 6.11 and 14.4, and 14.9).

So we now investigate complex series in general and series in powers of $z$ (or powers of $z - a$) in particular.

## Complex series

**6.1 Series of complex terms.** Suppose $\{a_n\}_{n\geqslant 0}$ is a complex sequence. The series $\sum a_n$ is said to **converge to the sum** $s$ if the sequence $\{s_n\}$ of partial sums, given by

$$s_n := a_0 + \cdots + a_n,$$

converges to the limit $s$, in the sense of the definition in 3.15. We write $s = \sum_{n=0}^{\infty} a_n$ (and this defines the expression on the right-hand side). Henceforth, where it would be pedantic to do otherwise, we do not distinguish between a convergent series $\sum a_n$ and the sum, $\sum_{n=0}^{\infty} a_n$, to which it converges. Parallel to 3.18(1), and following from it, we have the result that $\sum a_n$ converges if and only if the real series $\sum \operatorname{Re} a_n$ and $\sum \operatorname{Im} a_n$ both converge. Developing the theory of complex series is mainly a matter of checking that the same techniques work as in the real case.

We collect together for reference some basic facts about complex series. The corresponding results for the real case can be found, for example, in [**3**].

(1) **The terms of a convergent series** Suppose that $\sum a_n$ converges. Then

 (i) $a_n \to 0$ as $n \to \infty$, and

 (ii) there exists a real constant $M$ such that $|a_n| \leqslant M$ for all $n$.

(2) **Linear combinations of series** Suppose that $\sum a_n$ and $\sum b_n$ are convergent complex series. Then $\sum(a_n + kb_n)$ converges for any $k \in \mathbb{C}$ and

$$\sum_{n=0}^{\infty}(a_n + kb_n) = \sum_{n=0}^{\infty} a_n + k\sum_{n=0}^{\infty} b_n.$$

(3) **Absolute convergence vs. convergence** Suppose that the (real) series $\sum |a_n|$ converges. Then $\sum a_n$ converges. This result, which is expressed in words as 'absolute convergence implies convergence', can be obtained from its real analogue by considering real and imaginary parts. Alternatively, but essentially equivalently, it can be derived from the Cauchy criterion for convergence of a complex sequence, 3.21.

(4) **Testing for convergence** Let $\sum a_n$ be a complex series. The associated series $\sum |a_n|$ has real, non-negative terms. Hence well-known tests for convergence of series with non-negative terms can be applied. Combining this with (3) we get sufficient conditions for convergence of $\sum a_n$. Of particular importance for us are the following:

- **Comparison test** Suppose that $\sum b_n$ is a convergent series with $b_n \geqslant 0$ for all $n$ and suppose that, for some constant $k > 0$, $|a_n| \leqslant kb_n$ for all $n$. Then $\sum a_n$ converges absolutely, and hence converges.
- **d'Alembert's Ratio test** Assume that $\{a_n\}$ is such that

$$\ell := \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$$

exists. If $\ell < 1$ then $\sum |a_n|$ converges (and so $\sum a_n$ converges too). If $\ell > 1$ then $\sum |a_n|$ diverges. If $\ell = 1$ then the test gives no information.

Also worth noting is

- **Cauchy's $n$th-root test** Assume that $\{a_n\}$ is such that $\ell := \lim \sqrt[n]{|a_n|}$ exists. If $\ell < 1$ then $\sum |a_n|$ converges (and so $\sum a_n$ converges too). If $\ell > 1$ then $\sum |a_n|$ diverges. If $\ell = 1$ then the test gives no information.

**6.2 Geometric series.** We investigate the fundamental **geometric series** $\sum z^n$. We have already in 1.7 exploited the geometric identity

$$(1 - z)(1 + z + \cdots + z^n) = 1 - z^{n+1}.$$

This gives

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

We know that $\{z^{n+1}\}$ converges to 0 if $|z| < 1$, so $\sum z^n$ converges in this case. If $|z| \geqslant 1$ then the individual terms do not tend to 0, so the series diverges, by 6.1(1). We thus have the important result that

$$\sum z^n \quad \begin{cases} \text{converges, and } \sum_{n=0}^{\infty} z^n = (1-z)^{-1}, & \text{if } |z| < 1, \\ \text{fails to converge} & \text{if } |z| \geqslant 1. \end{cases}$$

We remark that we would be arguing in a circular fashion if we applied the tests in 6.1(4) to the geometric series: the proofs that validate these tests rely on knowledge of the behaviour of $\sum z^n$.

**6.3 Expansions derived from the geometric series.** The result in 6.2 can be viewed in two ways: either as summing an infinite series or as expanding $(1-z)^{-1}$ as a series when $|z| < 1$. Taking the second viewpoint, we may derive many related expansions. Here is a sample.

- $$\frac{1}{1+z} = 1 - z + z^2 - \cdots = \sum_{n=0}^{\infty} (-1)^n z^n \qquad (|z| < 1).$$
- $$\frac{1}{1-z^2} = 1 + z^2 + z^4 + \cdots = \sum_{n=0}^{\infty} z^{2n} \qquad (|z| < 1).$$
- $$\frac{1}{z-4} = -\frac{1}{4}\frac{1}{(1-(z/4))} = -\sum_{n=0}^{\infty} \frac{z^n}{4^{n+1}} \qquad (|z| < 4).$$

More generally, for $a, b \neq 0$,

- $$\frac{1}{az+b} = \frac{1}{b}\frac{1}{(1+(a/b)z)} = \sum_{n=0}^{\infty} (-1)^n \frac{a^n}{b^{n+1}} z^n \qquad (|z| < |b| \, / \, |a|).$$

Each of the expansions above is obtained by making a suitable change of variable in the standard geometric series. Other devices yield further expansions.

- By partial fractions: for $a \neq b$, we have

$$\begin{aligned} \frac{1}{(z-a)(z-b)} &= \frac{1}{(a-b)}\left(\frac{1}{z-a} - \frac{1}{z-b}\right) \\ &= \frac{1}{(a-b)}\left(\sum_{n=0}^{\infty}\left(\frac{1}{b^{n+1}} - \frac{1}{a^{n+1}}\right) z^n\right) \quad (|z| < \min\{|a|, |b|\}). \end{aligned}$$

- $$\frac{1}{1+z+z^2} = \frac{1-z}{1-z^3} = \sum_{n=0}^{\infty} \left(z^{3n} - z^{3n+1}\right) \qquad (|z| < 1).$$

The geometric series $\sum z^n$ is an expansion of $(1-z)^{-1}$ valid for $|z|<1$, that is, in the disc $\mathrm{D}(0;1)$. Frequently we want an expansion valid in a disc centred on some point $a \neq 0$. The following examples illustrate how to handle a change of origin.

- To obtain a series expansion of $(1-z)^{-1}$ valid in a disc centre $-3$, we may write

$$\frac{1}{1-z} = \frac{1}{4-(z+3)} = \sum_{n=0}^{\infty} \frac{1}{4^{n+1}}(z+3)^n \qquad (|z+3|<4).$$

More generally:

- $$\frac{1}{b-z} = \frac{1}{(b-a)-(z-a)} = \sum_{n=0}^{\infty} \frac{1}{(b-a)^{n+1}}(z-a)^n \quad (|z-a|<|b-a|).$$

## Power series

In the preceding section we exhibited expansions of various rational functions in powers of $z$ or, more generally, powers of $z-a$. We now turn things around and systematically study such series, called **power series**.

**6.4 Definition (power series and radius of convergence).** A **power series** is defined to be a series of the form $\sum c_n(z-a)^n$, where $a \in \mathbb{C}$ and $c_n \in \mathbb{C}$ $(n \geqslant 0)$. We shall henceforth often assume, without loss of generality, that $a=0$. Recall the difference between a power series and a polynomial: a polynomial has only *finitely many* terms.

The **radius of convergence** of the power series $\sum_{n=0}^{\infty} c_n(z-a)^n$ is defined to be

$$R := \sup\{\, |z| \, : \, \sum |c_n(z-a)^n| \text{ converges}\,\};$$

here we write $R=\infty$ if $\sum |c_n(z-a)^n|$ converges for arbitrarily large $|z-a|$. We have opted to define $R$ in terms of absolute convergence, rather than convergence: note that $\sum |c_n(z-a)^n|$ is a series of non-negative terms to which many convergence tests apply directly (recall 6.1(4)).

**6.5 Examples (calculating radius of convergence).**

- Consider $\sum nz^n$. We apply the Ratio test to $\sum |nz^n|$. For $z \neq 0$,

$$\left|\frac{(n+1)z^{n+1}}{nz^n}\right| = (1+1/n)\,|z| \to |z| \qquad \text{as } n \to \infty.$$

Hence $\sum |nz^n|$ converges if $|z| < 1$ and fails to converge if $|z| > 1$. We conclude that $R = 1$.

- Consider $\sum z^n/n!$. We apply the Ratio test to $\sum |z^n/n!|$. For $z \neq 0$,

$$\left|\frac{z^{n+1}/(n+1)!}{z^n/n!}\right| = \frac{|z|}{n+1} \to 0 \qquad \text{as } n \to \infty.$$

Hence $\sum |z^n/n!|$ converges for all $z$. We deduce that $R = \infty$.

- Consider $\sum n^n z^n$. The form of the series makes Cauchy's $n$th-root test a good choice:

$$\sqrt[n]{|n^n z^n|} = n\,|z| \to \begin{cases} 0 & \text{if } z = 0, \\ \infty & \text{otherwise.} \end{cases}$$

Therefore $\sum |n^n z^n|$ converges only for $z = 0$, so $R = 0$.

- Consider $\sum c_n z^n$, where

$$c_n = \begin{cases} m & \text{if } n = 2^m \text{ (for some } m = 0, 1, \ldots), \\ 0 & \text{otherwise.} \end{cases}$$

We cannot apply the Ratio test directly to $\sum |c_n z^n|$ because some of the terms are zero. But we can apply it to $\sum_{m=1}^{\infty} \left|mz^{2^m}\right|$. We have, for $z \neq 0$.

$$\frac{\left|(m+1)z^{2^{m+1}}\right|}{\left|mz^{2^m}\right|} = \frac{m+1}{m}\left|z^{(2^{m+1}-2^m)}\right| = \frac{m+1}{m}\,|z|^{2^m}$$
$$\to \begin{cases} 0 & \text{if } |z| < 1, \\ \infty & \text{if } |z| > 1. \end{cases}$$

We deduce that $R = 1$.

It is no coincidence that in each of the above examples the series converges absolutely for *all* $z$ such that $|z| < R$. The following lemma implies that every power series $f(z) = \sum c_n (z-a)^n$ with radius of convergence $R > 0$ has a 'disc of convergence', $\mathrm{D}(a; R)$. We prove later that $f$ is holomorphic in this disc. The series diverges for $|z - a| > R$. Any behaviour is possible on $|z - a| = R$: the series may always converge, may always diverge, or may converge at some points and diverge at others.

**6.6 Radius of convergence lemma.** Let $\sum c_n z^n$ be a power series with radius of convergence $R$.

(1) $\sum c_n z^n$ converges absolutely for all $z$ with $|z| < R$.

(2) $\sum c_n z^n$ fails to converge for any $z$ with $|z| > R$.

**Proof** (1) Let $|z| < R$. Then $R - \varepsilon$, where $\varepsilon := R - |z|$, is strictly less than $R$. Hence, by definition of supremum, there is some $w$ with $|z| = R - \varepsilon < |w| \leqslant R$ for which $\sum |c_n w^n|$ converges. Then $|c_n z^n| \leqslant |c_n w^n|$ for all $n$, so that $\sum |c_n z^n|$ converges by the Comparison test (6.1).

(2) Suppose, for a contradiction, that there is some $z$ with $|z| > R$ for which $\sum c_n z^n$ converges. Then, by 6.1(1), there exists $M$ such that $|c_n z^n| \leqslant M$ for all $n$. Pick $w$ such that $R < |w| < |z|$. Then

$$|c_n w^n| = |c_n z^n| \left|\frac{w^n}{z^n}\right| \leqslant M \left|\frac{w}{z}\right|^n .$$

The geometric series $\sum |w/z|^n$ converges, because $|w/z| < 1$. Hence, by the Comparison test, $\sum |c_n w^n|$ converges. This contradicts the definition of $R$. □

**6.7 Differentiating power series: preliminary remarks.** We have already hinted that there is a good connection between power series and holomorphic functions. How should we establish that $f(z) = \sum c_n z^n$ is differentiable for $|z| < R$, the radius of convergence? There is an obvious candidate for the derivative, namely $\sum n c_n z^{n-1}$. But this assumes that we can differentiate the series 'term-by-term'. Note that

$$\sum_{n=0}^{\infty} n c_n z^{n-1} = \sum_{n=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}z} c_n z^n$$

whereas

$$f'(z) = \frac{\mathrm{d}}{\mathrm{d}z}\left(\sum_{n-0}^{\infty} c_n z^n\right) :$$

the differentiation and summation are performed in different orders here. Both operations are performed by taking a limit. In general, limiting processes need not commute with one another. So it is *not* immediate that term-by-term differentiation of a power series is valid. Indeed, without proving it, we do not even know that $\sum n c_n z^{n-1}$ converges for $|z| < R$. It is more important to appreciate the need for justification of these statements than to master the technical details of their proofs, and we relegate the proofs to an optional appendix to this chapter.

Note that both (2) and, by induction, (4) in Theorem 6.8 are immediate from (3), so only the first and third statements need proving. A direct proof of (2) can be obtained from the results on uniform convergence in Chapter 14. An analogous theorem can, of course, be obtained for series of the form $\sum c_n (z-a)^n$.

**6.8 Differentiation theorem for power series.** Let $\sum c_n z^n$ have radius of convergence $R > 0$ and define $f$ in $\mathrm{D}(0;R)$ by $f(z) = \sum_{n=0}^{\infty} c_n z^n$. Then the following statements are true.

(1) $\sum n c_n z^{n-1}$ has radius of convergence $R$.

(2) $f$ is continuous in $\mathrm{D}(0;R)$.

(3) $f$ is holomorphic in $\mathrm{D}(0;R)$ and $f'$ is given by term-by-term differentiation:

$$f'(z) = \sum_{n=1}^{\infty} n c_n z^{n-1} \quad (|z| < R).$$

(4) $f$ has derivatives of all orders in $\mathrm{D}(0;R)$. Furthermore, $f^{(n)}(0) = n! c_n$ for $n \geqslant 0$.

**6.9 Example (exploiting the geometric series).** The geometric series $\sum z^n$ has radius of convergence 1, and provides a power series expansion of $(1-z)^{-1}$ for $|z| < 1$. By the Differentiation theorem,

$$(1-z)^{-2} = \frac{\mathrm{d}}{\mathrm{d}z}(1-z)^{-1} = 1 + 2z + 3z^2 + \ldots \quad (|z| < 1).$$

By induction we may obtain the binomial expansion of $(1-z)^{-n}$ $(n = 3, 4, \ldots)$. It has the same form as in the real case.

## A proof of the Differentiation theorem for power series

**6.10 Lemma.** The power series $\sum c_n z^n$ and $\sum n c_n z^{n-1}$ have the same radius of convergence.

**Proof** We first prove that, if $\sum |c_n z^n|$ converges for $|z| < R$, then $\sum \left| n c_n z^{n-1} \right|$ also converges for $|z| < R$. Choose $\rho$ such that $|z| < \rho < R$ and assume $z \neq 0$. Then

$$\left| n c_n z^{n-1} \right| = \frac{n}{|z|} \left( \frac{|z|}{\rho} \right)^n |c_n \rho|^n .$$

Since $|z|/\rho < 1$, the series $\sum n(|z|/\rho)^n$ converges by the Ratio test (see 6.5(1)). By 6.1(1), there exists a constant $M$ such that $n(|z|/\rho)^n \leqslant M$ for all $n$. Hence

$$\left| n c_n z^{n-1} \right| \leqslant \frac{M}{|z|} |c_n \rho^n| .$$

The result now follows from the Comparison test; see 6.1(4).

Conversely, suppose that $\sum \left|nc_n z^{n-1}\right|$ converges. Then

$$|c_n z^n| \leqslant |z| \left|nc_n z^{n-1}\right| \quad (n \geqslant 1),$$

so $\sum |c_n z^n|$ converges by the Comparison test. □

**6.11 Differentiation of power series.** Let $f(z) = \sum_{n=0}^{\infty} c_n z^n$ and assume that this power series has radius of convergence $R > 0$. Then $f \in \mathrm{H}(\mathrm{D}(0;R))$ and

$$f'(z) = \sum_{n=1}^{\infty} nc_n z^{n-1} \quad (|z| < R).$$

**Proof** Lemma 6.10 allows us to define

$$g(z) := \sum_{n=1}^{\infty} nc_n z^{n-1} \quad (|z| < R).$$

We want to show that $f'(z)$ exists and equals $g(z)$ for $z \in \mathrm{D}(0;R)$. For $z, z+h \in \mathrm{D}(0;R)$,

$$\frac{f(z+h) - f(z)}{h} - g(z) = \sum_{n=1}^{\infty} \left( \frac{(z+h)^n - z^n}{h} - nz^{n-1} \right),$$

and we must prove that this tends to 0 as $h \to 0$. We do this by estimating the terms in the series on the right-hand side. We shall need the binomial expansion

$$(z+h)^n = \sum_{k=0}^{n} \binom{n}{k} z^{n-k} h^k \qquad (n = 2, 3, \dots).$$

This expansion, valid for all $z$ and $h$ in $\mathbb{C}$, is proved by induction, just as the real version is: it relies solely on the arithmetic properties of $\mathbb{C}$ which mimic exactly those of $\mathbb{R}$.

So (notice the way in which the terms involving $z^n$ and $z^{n-1}$ cancel)

$$\begin{aligned}
\frac{(z+h)^n - z^n}{h} - nz^{n-1} &= \frac{nhz^{n-1} + \cdots + \binom{n}{k} h^k z^{n-k} + \cdots + h^n}{h} - nz^{n-1} \\
&= h\left( \binom{n}{2} z^{n-2} + \cdots + \binom{n}{k} h^{k-2} z^{n-k} + \cdots + h^{n-2} \right) \\
&= h \sum_{k=2}^{n} \binom{n}{k} h^{k-2} z^{n-k} \\
&= h \sum_{r=0}^{n-2} \frac{n!}{(n-(r+2))!(r+2)!} h^r z^{n-2-r}
\end{aligned}$$

(writing $r = k - 2$ for the last step). Hence, invoking the infinite version of the triangle inequality (Exercise 6.1),

$$\begin{aligned}\left|\frac{f(z+h)-f(z)}{h} - g(z)\right| &= \left|\sum_{n=1}^{\infty} c_n \left(h \sum_{r=0}^{n-2} \frac{n!}{(n-r-2)!(r+2)!} h^r z^{n-2-r}\right)\right| \\ &\leqslant |h| \sum_{r=0}^{n-2} |c_n| \frac{n!}{(n-r-2)!(r+2)!} |h|^r |z|^{n-2-r} \\ &\leqslant |h| \sum_{n=1}^{\infty} n(n-1)|c_n| \left(\sum_{r=0}^{n-2} \frac{(n-2)!}{(n-2-r)!r!} |h|^r |z|^{n-2-r}\right) \\ &= |h| \sum_{n=1}^{\infty} n(n-1)|c_n| (|z|+|h|)^{n-2}.\end{aligned}$$

Fix $z$ and choose $\rho$ with $|z| < \rho < R$, so that $|z|+|h| < \rho$ whenever $|h| < \rho - |z|$. By Lemma 6.10, used twice over, $\sum_{n=2}^{\infty} n(n-1)|c_n| \rho^{n-2}$ converges, to a finite constant independent of $h$. We conclude that $f'(z)$ does indeed exist and equal $g(z)$. □

## Exercises

6.1 Prove that, if $a_n$ $(n = 0, 1, \dots)$ are complex numbers such that $\sum |a_n|$ converges, then

$$\left|\sum_{n=0}^{\infty} a_n\right| \leqslant \sum_{n=0}^{\infty} |a_n|.$$

6.2 Write down an expansion of the form $\sum_{n=0}^{\infty} c_n z^n$ for

(i) $\dfrac{1}{2z+5}$, (ii) $\dfrac{1}{1+z^4}$, (iii) $\dfrac{1+\mathrm{i}z}{1-\mathrm{i}z}$,

(iv) $\dfrac{1}{1-z+z^2}$, (v) $\dfrac{1}{(z+1)(z+2)}$, (vi) $\dfrac{1}{(z^2-1)(z^2-9)}$.

In each case, specify where the expansion is valid.

6.3 Write down an expansion of (i) $(1-z)^{-1}$ and (ii) $1/(z(z+2))$ as a power series in powers of (a) $z+1$ and (b) $z - \mathrm{i}$.

6.4 For each of the following power series, calculate the radius of convergence and hence find at which points the series defines a holomorphic function.

(i) $\sum_{n=1}^{\infty}(-1)^n z^n/n^3$, (ii) $\sum_{n=0}^{\infty} z^{5n}$,

(iii) $\sum_{n=1}^{\infty} z^n/n^n$, (iv) $\sum_{n=0}^{\infty} n!z^n$.

6.5 Let $\sum c_n z^n$ have radius of convergence $R$. Prove that

$$R = \sup\{\,|z| \,:\, \sum c_n z^n \text{ converges}\} = \sup\{\,|z| \,:\, c_n z^n \to 0 \text{ as } n \to \infty\,\}.$$

(Hint: look at the proof of the Radius of convergence lemma (6.6).)

6.6 let $f(z) = \sum_{n=0}^{\infty} c_n z^n$ have radius of convergence $R$.

(i) Prove that $\sum_{n=0}^{\infty} \overline{c_n} z^n$ also has radius of convergence $R$.

(ii) Prove that $\overline{f(\overline{z})} = \sum_{n=0}^{\infty} \overline{c_n} z^n$ for $|z| < R$. (Hint: exploit the fact that $z \mapsto \overline{z}$ is a continuous function.)

(iii) Deduce that $g$, defined by $g(z) = \overline{f(\overline{z})}$ is holomorphic in $\mathrm{D}(0; R)$. (cf. Exercise 5.11.)

6.7 Obtain power series expansions for $(1+z)^{-2}$ and for $(1+z)^{-3}$, each valid for $|z| < 1$. (Hint: use 6.11.)

6.8 Let $p$ be a polynomial of degree $k > 0$. Prove that $\sum p(n)z^n$ has radius of convergence 1 and that there exists a polynomial $q(z)$ of degree $k$ such that

$$\sum_{n=0}^{\infty} p(n)z^n = q(z)(1-z)^{-(k+1)} \quad (|z| < 1).$$

6.9 let $p(z) = (z-\alpha_1)(z-\alpha_2)\ldots(z-\alpha_N)$, where $\alpha_1, \alpha_2, \ldots, \alpha_N$ are distinct complex numbers. Let $M = \min_{1 \leqslant k \leqslant N} |\alpha_k|$. Prove that it is possible to represent $1/p(z)$ as a power series $\sum_{n=0}^{\infty} c_n z^n$, for $|z| < M$. Could the radius of convergence of this power series exceed $M$?

6.10 Determine for which values of $z$ the following series converge absolutely:

(i) $\sum_{n=0}^{\infty} \dfrac{(z+1)^n}{2^n}$, (ii) $\sum_{n=0}^{\infty} \left(\dfrac{z-1}{z+1}\right)^n$,

(iii) $\sum_{n=1}^{\infty} \dfrac{1}{n^2}(z^n + z^{-n})$, (iv) $\sum_{n=0}^{\infty} \dfrac{z^n}{1-z^n}$.

# 7 A cornucopia of holomorphic functions

We shall assume that readers know the basic properties of the real trigonometric, exponential, and hyperbolic functions, including the formulae for their derivatives from which the Maclaurin expansions of these functions are obtained. For most, this knowledge will be founded on a naive treatment of these functions, relying on elementary geometry and trigonometry. A few may have seen the naive approach superseded by a more analytical one, in which functions are *defined* by power series and their expected properties are then derived from the properties of these series. In the complex case, a geometric approach to the elementary functions is no longer available, but this is not a problem. Power series definitions serve admirably, since convergent complex power series behave so well.

We begin by investigating that most fundamental of functions, the exponential function.

## The exponential function

**7.1 Definition (exponential function).** We have already shown in 6.5 that the power series $\sum z^n/n!$ has infinite radius of convergence. We may therefore define the exponential function by

$$\mathrm{e}^z := \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad (z \in \mathbb{C}).$$

(Note that we do not yet know that this is compatible with our earlier use of the symbol $\mathrm{e}^{\mathrm{i}\theta}$ to denote $\cos\theta + \mathrm{i}\sin\theta$ for $\theta \in \mathbb{R}$. The reconciliation comes in 7.6, after we have investigated trigonometric functions.)

We may immediately record what Theorem 6.11 tells us about the exponential function.

**7.2 Theorem (holomorphy of the exponential function).** The function $e^z$ is holomorphic (and hence also continuous) in $\mathbb{C}$ and

$$\frac{d}{dz}e^z = e^z \quad \text{for all } z \in \mathbb{C}.$$

There is a useful technique for deriving certain functional identities which relies on the fact that, in an open disc, a holomorphic function with zero derivative is necessarily constant. The proof of our next result about complex exponentials illustrates this.

**7.3 Theorem (properties of exponentials).**

(1) $e^0 = 1$;

(2) $e^{z+w} = e^z e^w$ for all $z, w \in \mathbb{C}$;

(3) $e^z \neq 0$ for all $z \in \mathbb{C}$.

**Proof** (1) is immediate from the series definition.

For (2), we fix $c \in \mathbb{C}$ and consider

$$f(z) := e^z e^{c-z}.$$

By 7.2 and the Chain rule (5.8(2)),

$$f'(z) = e^z e^{c-z} - e^z e^{c-z} = 0.$$

Therefore, by 5.12, there exists a constant $K$ such that $f(z) = K$ for all $z \in \mathbb{C}$; here $K$ depends on $c$. To find $K$, we put $z = c$ and obtain $K = e^c e^{c-c}$. So $K = e^c$, by (1). Thus $e^c = e^z e^{c-z}$ for all $z, c \in \mathbb{C}$. Choosing $c = w - z$ we get (2).

From (1) and (2) we have $e^z e^{-z} = 1$, so (3) follows. □

**7.4 The modulus of an exponential.** Let $z = x + iy$, where $x, y \in \mathbb{R}$. Then $|e^z| = e^x$. In particular, $\left|e^{iy}\right| = 1$ for all $y \in \mathbb{R}$.

To prove the first assertion, observe that

$$\begin{aligned}
|e^z|^2 &= e^z\overline{e^z} && \text{(by 1.8(3))}\\
&= e^z e^{\overline{z}} && \text{(by Exercise 6.6)}\\
&= e^{z+\overline{z}} && \text{(by 7.3)}\\
&= e^{2x} && \text{(by 1.8(4))}\\
&= (e^x)^2 && \text{(by 7.3).}
\end{aligned}$$

Hence $|e^z| = e^x$ (since both sides are real and positive). The second assertion comes from taking $x = 0$. □

## Complex trigonometric and hyperbolic functions

We shall use power series to define the functions $\cos z$ and $\sin z$ and their hyperbolic analogues $\cosh z$ and $\sinh z$. We shall see that, on the $x$- and $y$-axes, $\cos z$ behaves, respectively, like a real cosine and like a real cosh function, and as a hybrid between these two elsewhere (and similarly for $\sin z$).

**7.5 Definitions (trigonometric and hyperbolic functions).** We define, for $z \in \mathbb{C}$,

$$\cos z := 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!},$$

$$\cosh z := 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!},$$

$$\sin z := z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!},$$

$$\sinh z := z + \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!}.$$

The Ratio test shows, easily, that all four series have infinite radius of convergence and so these series do indeed define functions with domain $\mathbb{C}$.

Theorem 6.11 tells us that the four functions are holomorphic in $\mathbb{C}$ and allows us to calculate their derivatives by differentiating term-by-term. No surprises:

$$\frac{d}{dz} \cos z = -\sin z, \qquad \frac{d}{dz} \sin z = \cos z,$$
$$\frac{d}{dz} \cosh z = \sinh z, \qquad \frac{d}{dz} \sinh z = \cosh z.$$

**7.6 Key relationships.** Immediately from the definitions (and 6.1(2)) we have

$$e^{iz} = \cos z + i \sin z \quad (\text{for } z \in \mathbb{C}).$$

In particular, as previously promised, we obtain

$$e^{i\theta} = \cos\theta + i\sin\theta \quad (\text{for } \theta \in \mathbb{R}).$$

Note that $|e^{i\theta}| = 1$ emerges from this if we assume, as we did in 1.3, properties of cosine and sine on $\mathbb{R}$. We did not need such assumptions in the proof in 7.4.

In the other direction, we can express the trigonometric functions in terms of exponentials: for any $z \in \mathbb{C}$,

$$\cos z = \frac{1}{2}(e^{iz} + e^{-iz}) \quad \text{and} \quad \sin z = \frac{1}{2i}(e^{iz} - e^{-iz}).$$

Similarly,

$$\cosh z = \frac{1}{2}(e^{z} + e^{-z}) \quad \text{and} \quad \sinh z = \frac{1}{2}(e^{z} - e^{-z}).$$

Comparing these formulae, or working directly with the defining series, we obtain the relationships

$$\cos iz = \cosh z \quad \text{and} \quad \sin iz = i\sinh z$$

(known as **Osborn's rules**).

**7.7 Addition formulae.** Nothing novel here! Exactly the same addition formulae hold for the complex trigonometric and hyperbolic functions as for their real counterparts. For example,

$$\cos(z + w) = \cos z \cos w - \sin z \sin w.$$

To prove this, we use 7.6 and 7.3:

$$\begin{aligned}\cos z \cos w - &\sin z \sin w \\ &= \frac{1}{4}\big((e^{iz} + e^{-iz})(e^{iw} + e^{-iw}) + (e^{iz} - e^{-iz})(e^{iw} - e^{-iw})\big) \\ &= \frac{1}{2}\big(e^{iz}e^{iw} + e^{-iz}e^{-iw}\big) \\ &= \cos(z + w).\end{aligned}$$

Later (in 15.11) we shall see how identities between holomorphic functions $f$ and $g$ which are true on $\mathbb{R}$ (or on a non-empty open subinterval of $\mathbb{R}$) persist in the intersection of the sets where $f$ and $g$ are holomorphic (subject to some topological qualifications).

**7.8 Real and imaginary parts.** We can apply the addition formulae and Osborn's rules to obtain, for $z = x + \mathrm{i}y$,

$$\begin{aligned} e^{x+iy} &= e^x e^{iy} = e^x(\cos y + i\sin y), \\ \cos(x+iy) &= \cos x\cos iy - \sin x\sin iy = \cos x\cosh y - i\sin x\sinh y, \\ \sin(x+iy) &= \sin x\cos iy + \cos x\sin iy = \sin x\cosh y + i\cos x\sinh y. \end{aligned}$$

These formulae will be used in the next section to identify the points where the trigonometric functions, and likewise the hyperbolic functions, take the value zero—necessary if we are to define $\operatorname{cosec} z = 1/\sin z$, $\tan z = \sin z/\cos z$, and so on. They also show how each of the complex trigonometric and hyperbolic functions can be seen as a hybrid of real trigonometric and hyperbolic functions.

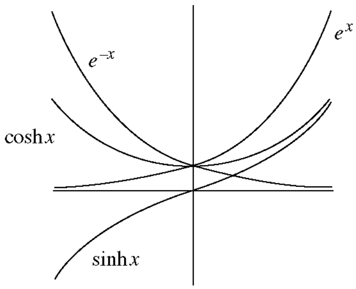



**Figure 7.1** The real hyperbolic functions

**7.9 Unboundedness.** We can prove from the series definitions that $\cosh x$ and $\sinh x$ tend to infinity as the real variable $x \to \infty$. As a reminder of the way these real hyperbolic functions behave, their graphs are shown in Fig. 7.1. From Osborn's rules we have

$$\begin{aligned} |\cos iy| &= |\cosh y| \to \infty \quad \text{as } y \to \infty, \\ |\sin iy| &= |\sinh y| \to \infty \quad \text{as } y \to \infty. \end{aligned}$$

This is in stark contrast to the behaviour of the real functions:

$$|\cos x| \leqslant 1 \quad \text{and} \quad |\sin x| \leqslant 1 \qquad (x \in \mathbb{R}).$$

Not only are these inequalities no longer valid when $x \in \mathbb{R}$ is replaced by $z \in \mathbb{C}$ but the functions $\cos z$ and $\sin z$ are *unbounded* in $\mathbb{C}$. For further information see Exercise 7.6.

## Zeros and periodicity

In this section we begin to explore the far-reaching consequences for complex analysis of the fundamental equation

$$e^{2\pi i} = 1.$$

This equation comes from the relation $e^{i\theta} = \cos\theta + i\sin\theta$, assuming that $\sin 2\pi = 0$ and $\cos 2\pi = 1$. [Purists who seek an ultra-formal approach, independent of elementary trigonometry, may *define* $\pi$ to be the smallest positive solution of the equation $\sin\theta = 0$, where $\sin\theta$ is defined by its power series.]

We know that if $f$ is holomorphic in some open set $G$ then $1/f$ is holomorphic in $G$ provided $f(z) \neq 0$ for $z \in G$. It is therefore essential to identify where frequently occurring functions take the value zero. We have already (in 1.7) looked at the zeros (roots) of certain polynomials. We now do the same for the elementary functions.

**7.10 Solving equations, finding zeros.** We cannot emphasize too strongly the importance of the following facts:

$$e^z = 1 \iff z = 2k\pi i \quad (k \in \mathbb{Z}),$$
$$e^z = -1 \iff z = (2k+1)\pi i \quad (k \in \mathbb{Z}).$$

To prove the first of these claims we write $z = x + iy$ and solve the equation $e^x(\cos y + i\sin y) = 1$. Taking the modulus of both sides gives $e^x = 1$, and this holds if and only if $x = 0$. As in 1.3, we then require $\cos y = 1$ and $\sin y = 0$, and this is satisfied if and only if $y = 2k\pi$ for some $k \in \mathbb{Z}$. The solutions of $e^z = -1$ are obtained in a similar way.

We have already seen (7.3) that $e^z \neq 0$ for any $z \in \mathbb{C}$. On the other hand, the functions $e^z - 1$ and $e^z + 1$ both take the value zero at infinitely many points. Contrast this with the real case: for $x$ *real*, $e^x - 1 = 0$ only if $x = 0$ and $e^x + 1$ is never zero.

For the trigonometric and hyperbolic functions we have

$$\cos z = 0 \iff z = \frac{1}{2}(2k+1)\pi \quad (k \in \mathbb{Z}),$$
$$\sin z = 0 \iff z = k\pi \quad (k \in \mathbb{Z}),$$
$$\cosh z = 0 \iff z = \frac{1}{2}(2k+1)\pi i \quad (k \in \mathbb{Z}),$$
$$\sinh z = 0 \iff z = k\pi i \quad (k \in \mathbb{Z}).$$

We draw attention particularly to the latter two claims, where the situation is quite different from that in the real case.

We can now see that $\tan z := \sin z/\cos z$ is defined, and is holomorphic, in $\mathbb{C} \setminus \{\, (2k+1)\pi/2 \,:\, k \in \mathbb{Z} \,\}$ and that $\cot z := \cos z/\sin z$, similarly, exists and is holomorphic except at the points $k\pi$ $(k \in \mathbb{Z})$.

**7.11 Example (solving equations).** We include an example of the solution of equations, to reinforce readers' appreciation of how different the complex-variable trigonometric functions are from their real-variable counterparts.

We shall find the solutions of the equation $\sin z = 10^3$. We have

$$\begin{aligned}\sin(x+\mathrm{i}y) = 10^3 &\iff \sin x \cosh y + \mathrm{i} \cos x \sinh y = 10^3 \\ &\iff \cos x \sinh y = 0 \text{ and } \sin x \cosh y = 10^3.\end{aligned}$$

From the former condition either (a) $y = 0$ or (b) $x = (2k+1)\pi/2$ $(k \in \mathbb{Z})$. In case (a), $\cosh y = 1$ and so we require $\sin x = 10^3$, which is impossible. Case (b) leads to no solutions if $k$ is odd, because $\cosh y > 0$; for $k$ even, we must have $\sin(2k+1)\pi/2 = 1$ and $\cosh y = 10^3$, that is, $y$ takes the unique value $\cosh^{-1}(10^3)$. Hence

$$\sin z = 10^3 \iff z = \tfrac{1}{2}(4m+1)\pi + \mathrm{i}\cosh^{-1}(10^3) \text{ for some } m \in \mathbb{Z}.$$

**7.12 Periodicity.** The results in 7.10 tell us that $\sin z$ and $\cos z$ are periodic of period $2\pi$, just as in the real case. Also

$$\mathrm{e}^{z+\alpha} = \mathrm{e}^{z} \text{ for all } z \in \mathbb{C} \iff \alpha = 2k\pi\mathrm{i} \text{ for some } k \in \mathbb{Z}.$$

Therefore $\mathrm{e}^z$ is periodic, of period $2\pi\mathrm{i}$. This periodicity, of course, stems from the equation $\mathrm{e}^{2\pi\mathrm{i}} = 1$.

The functions $\cos z$ and $\sin z$ are, as in the real case, periodic of period $2\pi$; $\cosh z$ and $\sinh z$ have periodicity $2\pi\mathrm{i}$.

## Argument, logarithms, and powers

We now start to come to terms with the many-valuedness of the complex analogues of some important real functions, namely logarithms and powers. Our approach will initially be static. We focus on values at a fixed, but arbitrary, point $z$ of the domain of a multifunction. Only later do we think dynamically and consider $z$ as a variable.

**7.13 Argument.** As we have seen, the periodicity of the exponential function has awkward consequences for the polar representation of complex numbers: the angle $\theta$ in the expression $z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$ is not uniquely determined. Indeed, this is the fundamental cause of many-valuedness in complex function theory.

For any $z \neq 0$, we define the **argument** of $z$ to be

$$[\![\arg z]\!] := \{\,\theta \in \mathbb{R} : z = |z|\,\mathrm{e}^{\mathrm{i}\theta}\,\}.$$

The bracket notation $[\![\arg z]\!]$ is designed to emphasize that the argument of $z$ is a *set* of numbers, not a single number. In fact, $[\![\arg z]\!]$ is an infinite set, consisting of all numbers of the form $\theta + 2k\pi$ for $k \in \mathbb{Z}$, where $\theta$ is any fixed real number such that $\mathrm{e}^{\mathrm{i}\theta} = z/\,|z|$. For example, $[\![\arg \mathrm{i}]\!] = \{\,(4k+1)\pi/2 : k \in \mathbb{Z}\,\}$.

For $z, w \neq 0$,

$$\begin{aligned}
[\![\arg(zw)]\!] &= \{\,\theta + \varphi : \theta \in [\![\arg z]\!],\ \varphi \in [\![\arg w]\!]\,\},\\
[\![\arg(1/z)]\!] &= \{\,-\theta : \theta \in [\![\arg z]\!]\,\}.
\end{aligned}$$

**7.14 Complex logarithms: the inverse of an exponential.** We may define the logarithm on $(0,\infty) \subseteq \mathbb{R}$ as the function which is inverse to the exponential function: for each positive real number $x$, there exists a unique real solution $t = \log_{\mathrm{e}} x$ to the equation $\mathrm{e}^t = x$. (Since we shall work exclusively with logarithms to the base e, we shall henceforth drop the subscript and write $\log x$ in place of $\ln x = \log_{\mathrm{e}} x$.) In the complex case we seek solutions to the equation $\mathrm{e}^w = z$.

Suppose $z \in \mathbb{C}$, $z \neq 0$. Put $z = \mathrm{e}^w = \mathrm{e}^{u+\mathrm{i}v}$ $(u, v$ real). Then

$$|z| = \left|\mathrm{e}^u \mathrm{e}^{\mathrm{i}v}\right| = \mathrm{e}^u \qquad \text{(by 7.3)}$$

and

$$[\![\arg z]\!] = \{\,v + 2k\pi : k \in \mathbb{Z}\,\} \qquad \text{(see 7.10).}$$

We have derived the important relation

$$\mathrm{e}^w = z \iff w = \log|z| + \mathrm{i}\theta, \quad \text{where } \theta \in [\![\arg z]\!].$$

We accordingly *define*, for $z \neq 0$,

$$[\![\log z]\!] := \{\,\log|z| + \mathrm{i}\theta : \theta \in [\![\arg z]\!]\,\}.$$

For example,

$$\llbracket \log 2 \rrbracket = \{ \log 2 + 2k\pi \mathrm{i} : k \in \mathbb{Z} \},$$
$$\llbracket \log(-1) \rrbracket = \{ (4k-1)\pi \mathrm{i}/2 : k \in \mathbb{Z} \}.$$

Also, for $w = u + \mathrm{i}v$,

$$\llbracket \log \mathrm{e}^w \rrbracket = \{ \log \mathrm{e}^u + \mathrm{i}(v + 2k\pi) : k \in \mathbb{Z} \} = \{ w + 2k\pi : k \in \mathbb{Z} \}.$$

By contrast,

$$\mathrm{e}^w = z \quad \text{for } \textit{any } w \in \llbracket \log z \rrbracket.$$

**7.15 Powers.** If $n$ is a positive integer and $z \neq 0$, there exist $n$ solutions to the equation $w^n = z$, given in terms of the polar representation $z = r\mathrm{e}^{\mathrm{i}\theta}$ by $w = r^{1/n}\mathrm{e}^{2k\pi \mathrm{i}/n}$ $(k = 0, \ldots, n-1)$; recall 1.7. More generally, if $\alpha$ is a complex number, we define, for $z \neq 0$,

$$\llbracket z^\alpha \rrbracket := \{ \mathrm{e}^{\alpha(\log|z| + \mathrm{i}\theta)} : \theta \in \llbracket \arg z \rrbracket \}.$$

The motivation for this definition comes from the real analogue: $a^x := \mathrm{e}^{x \log a}$ for $a > 0$. Note that $\mathrm{e}^\alpha$ (as defined in 7.3) is one member of $\llbracket \mathrm{e}^\alpha \rrbracket$.

Only when $\alpha$ is an integer $n$ does $\llbracket z^\alpha \rrbracket$ not produce multiple values: in this case $\llbracket z^n \rrbracket$ contains the single point $z^n$. When $\alpha = 1/n$ $(n = 2, 3, \ldots)$ $\llbracket z^\alpha \rrbracket$ contains the values of the $n$th root given above.

Complex powers must be treated with circumspection. The formula $x^\alpha x^\beta = x^{\alpha+\beta}$ $(x > 0,\ \alpha, \beta \text{ real})$ can be shown to have a complex analogue, in which the values of the multifunctions have to be appropriately selected. But $x_1^\alpha x_2^\alpha = (x_1 x_2)^\alpha$ $(x_1, x_2 > 0,\ \alpha \text{ real})$ has no universally valid complex generalization.

## Holomorphic branches of some simple multifunctions

Many-valuedness is not a serious problem so long as we are concerned with selecting a value of a multifunction at some *fixed* point $z$. But it is a very different matter when $z$ is allowed to roam freely in the plane. In this section we show how to extract holomorphic functions from $\llbracket \log z \rrbracket$ and $\llbracket z^\alpha \rrbracket$ $(\alpha \notin \mathbb{Z})$. Our treatment is economical, but suffices for working with these functions in succeeding chapters. We explore many-valuedness in greater depth in Chapter 9 and show how to handle more complicated multifunctions, in particular logarithms and powers of rational functions. Chapter 9 is optional and only Chapter 23) directly depends on it.

**7.16 Holomorphic branches of the logarithm.** Consider

$$[\![\log z]\!] = \{\log|z| + \mathrm{i}\theta : \theta \in [\![\arg z]\!]\} \quad (z \neq 0).$$

The many-valuedness arises because values of $\theta$ differing by an integer multiple of $2\pi$ give the same point $z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$ but give different values of $\log|z| + \mathrm{i}\theta$. We get a 1-valued function if we restrict $\theta$ either to $[0, 2\pi)$ or to $(-\pi, \pi]$ (that is, if we take a principal-value determination of the argument of $z$). These ranges are the most usual choices, but any interval of length $2\pi$, closed at one end and open at the other, serves equally well.

For definiteness, let us restrict $\theta$ to $(-\pi, \pi]$ and remind ourselves of this fact by putting a **cut** (also called a **branch cut**) in the plane along the non-positive real axis $(-\infty, 0]$ and forbidding $z$ to cross the cut. We regard the cut as having two edges: on the upper edge, $\theta = \pi$ and on the lower edge, $\theta = -\pi$. Our choice of restriction on $\theta$ means that we identify the upper edge of the cut with the points of $(-\infty, 0]$ and exclude the lower edge of the cut from our plane. (But see also 19.11 below, where we introduce a convenient modification of this convention.) The point 0 is called a **branch point**. See 9.3 for a general discussion of branch points.

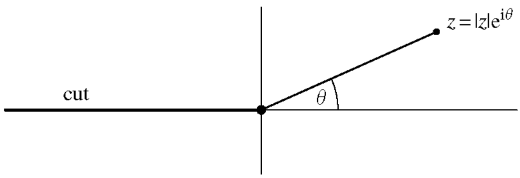



**Figure 7.2** Plane cut along $(-\infty, 0]$

In the cut plane, define

$$f_k(z) = \log r + \mathrm{i}(\theta + 2k\pi) \quad (0 \neq z = r\mathrm{e}^{\mathrm{i}\theta},\ -\pi < \theta \leqslant \pi),$$

for $k \in \mathbb{Z}$. Then $[\![\log z]\!] = \{f_k(z) : k \in \mathbb{Z}\}$. Each $f_k$ is 1-valued. Certainly $z \mapsto \operatorname{Re} f_k(z) = \log|z|$ is continuous for $z \neq 0$, by continuity of the real logarithm. Also $\operatorname{Im} f_k$ is continuous at any point not on the cut (see Exercise 7.17). So $f_k$ is continuous. However $f_k$ is discontinuous at points on the cut. In crossing the cut from the upper half-plane to the lower half-plane we transfer from $f_k$ to $f_{k+1}$. The transfer is continuous in the sense that $\lim_{h\to 0+} f_k(\mathrm{i}h) = \lim_{h \to 0-} f_{k+1}(\mathrm{i}h)$.

The function $f_k$ is holomorphic in $\mathbb{C}_\pi := \mathbb{C} \setminus (-\infty, 0]$, with $f_k'(z) = 1/z$. To see this, write $\zeta = f_k(z)$ and $\eta = f_k(z+h) - f_k(z)$. Continuity of $f_k$ in $\mathbb{C}_\pi$ implies that $\eta \to 0$ as $\zeta \to 0$. Hence

$$\frac{f_k(z+h) - f_k(z)}{h} = \frac{\eta}{e^{\zeta+\eta} - e^{\zeta}} \to \frac{1}{e^{\zeta}} = \frac{1}{z} \quad \text{as } h \to 0.$$

Here we have used the fact that $e^{f_k(z)} = z$ for any $z \in \mathbb{C}_\pi$ and properties of the exponential given in 7.2. We can alternatively verify that $f_k \in \mathrm{H}(\mathbb{C}_\pi)$ by checking that its real and imaginary parts have continuous first-order partial derivatives and satisfy the Cauchy–Riemann equations (see 5.6). We call the functions $f_k$ **holomorphic branches** of the logarithm.

**7.17 Holomorphic branches of powers.** Let $n \in \mathbb{Z}$, $n \neq 0, \pm 1$. Consider the multifunction

$$[\![ z^{1/n} ]\!] = \{ |z|^{1/n} e^{i\theta/n} : 0 \neq z = |z| e^{i\theta} \}.$$

We elect to restrict $\theta$ to $[0, 2\pi)$ and so cut the plane along the non-negative real axis. We define

$$g_k(z) := e^{2k\pi i/n} r^{1/n} e^{i\theta/n} \quad (0 \neq z = re^{i\theta}).$$

The functions $g_k$ are 1-valued and such that $(g_k(z))^n = z$ for each point in the cut plane. We have $g_k = g_{k\,(\mathrm{mod}\,n)}$ for any $k \in \mathbb{Z}$ and

$$[\![ z^{1/n} ]\!] = \{ g_k(z) : 0 \leqslant k \leqslant n-1 \}.$$

Each $g_k$ is continuous except at points of the cut. Arguments similar to those for the logarithm confirm that $g_k$ is holomorphic at any point of $\mathbb{C} \setminus [0, \infty)$. Each is called a holomorphic branch of the $n$th root.

Similar considerations apply to $[\![ z^{\alpha} ]\!]$ for an arbitrary $\alpha$, save that there will, in general, be infinitely many values at each point and infinitely many holomorphic branches.

## Exercises

**Exercises from the text.** Verify the claims about radius of convergence and derivatives in 7.5. Verify the unproved formulae given in 7.6 and 7.7. Prove (see 7.10) that $e^z = -1$ if and only if $z = (2k+1)\pi i$ $(k \in \mathbb{Z})$.

7.1 Find the real and imaginary parts of (i) $e^{2z}$, (ii) $e^{z^2}$, (iii) $e^{e^z}$.

7.2 Suppose $\alpha^3 = 1$, $\alpha \neq 1$. Express $e^z + e^{\alpha z} + e^{\alpha^2 z}$ as a power series. Hence evaluate $\sum_{n=0}^{\infty} 8^n/(3n)!$. Find also $\sum_{n=0}^{\infty} (27)^n/(3n+1)!$.

7.3 Use Exercise 6.1 to prove that

(i) for all $z \in \mathbb{C}$,

$$|e^z - 1| \leqslant e^{|z|} - 1 \leqslant |z|\, e^{|z|};$$

(ii) for all $z \in \overline{\mathbb{D}}(0;1)$,

$$(3 - e)\,|z| \leqslant |e^z - 1| \leqslant (e - 1)\,|z|\,.$$

7.4 Let $f$ be holomorphic in a region $G$. Let $g$ be defined by $g(z) = e^{f(z)}$. Prove that if $g$ is constant then $f$ is constant.

7.5 Prove that $\overline{\cos z} = \cos \overline{z}$ and $\overline{\sin z} = \sin \overline{z}$, for all $z \in \mathbb{C}$. Using Exercise 5.11 or otherwise, prove that $\overline{\cos z}$ and $\overline{\sin z}$ are not holomorphic at any point of $\mathbb{C}$.

7.6 Prove with the aid of 1.8(6) and the preceding exercise that

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

Deduce that

$$|\sinh y| \leqslant |\cos z| \leqslant \cosh y.$$

(Analogous results can obviously be proved for $\sin z$.)

7.7 Prove that $|\cos^2 z| + |\sin^2 z| = 1$ is false if $z = x + iy$ with $y \neq 0$.

7.8 Let $z = Re^{i\alpha}$, where $\alpha$ is fixed and $0 < \alpha < \pi/2$. Prove that each of (i) $|e^{-iz}|$, (ii) $|\cos z|$, (iii) $|\sinh z|$ tends to infinity as $R \to \infty$.
Describe the behaviour of these functions as $R \to \infty$ when (a) $\alpha = 0$ and (b) $\alpha = \pi/2$.

7.9 (a) Let $f(z) = e^{iz}$. Find all values of $z$ for which $f(z)$ is (i) real, (ii) purely imaginary, and (iii) of modulus $< 1$.

(b) Repeat (a) with $f(z) = e^{-2z}$.

7.10 Find all solutions to the following equations:

$$\text{(i)}\ \cosh z = -1, \qquad \text{(ii)}\ \cos^2 z = 4, \qquad \text{(iii)}\ \tan z = i.$$

7.11 Find $Z(f) := \{\, z \in \mathbb{C} : f(z) = 0 \,\}$ for each of the following functions $f$:

$$\text{(i)}\ (z^4 - 1)\sin \pi z, \qquad \text{(ii)}\ \cosh^2 z, \qquad \text{(iii)}\ 1 + e^{2z},$$
$$\text{(iv)}\ \sin^3(1/z)\ \ (z \neq 0), \qquad \text{(v)}\ 1 - e^{z^2}, \qquad \text{(vi)}\ 1 + e^{z^2}.$$

(In (v) and (vi), express the elements of $Z(f)$ in polar form.)

7.12 Define $f$ by

$$f(z) = \begin{cases} \dfrac{\sin z}{z} & \text{if } z \neq 0, \\ 1 & \text{if } z = 0. \end{cases}$$

By considering a power series expansion, prove that $f$ is holomorphic in $\mathbb{C}$.

7.13 (a) Find each of $[\![\arg(-1)]\!]$, $[\![\arg(1-\mathrm{i})]\!]$, $[\![\arg \mathrm{e}^{-2\pi\mathrm{i}/3}]\!]$.
(b) Find each of $[\![\log(-1)]\!]$, $[\![\log(1+\mathrm{i})]\!]$, $[\![\log\omega]\!]$.

7.14 Find each of $[\![1^{\frac{1}{3}}]\!]$, $[\![(-8)^{\frac{1}{3}}]\!]$, $[\![\mathrm{i}^{\frac{1}{6}}]\!]$, $[\![(-1)^{\frac{1}{4}}]\!]$. In each case, plot the set as a subset of the complex plane.

7.15 Find $[\![\sqrt{2}^{\,\mathrm{i}}]\!]$, $[\![\mathrm{i}^{\sqrt{2}}]\!]$, $[\![\mathrm{i}^{\mathrm{i}}]\!]$, and $[\![\mathrm{e}^{\mathrm{i}\pi}]\!]$.

7.16 (a) Show that, if $w \in [\![z^{\alpha+\beta}]\!]$, then there exist $z_1 \in [\![z^{\alpha}]\!]$ and $z_2 \in [\![z^{\beta}]\!]$ such that $w = z_1 z_2$.
(b) Give an example to show that it may happen that $z_1 \in [\![z^{\alpha}]\!]$ and $z_2 \in [\![z^{\beta}]\!]$, yet $z_1 z_2 \notin [\![z^{\alpha+\beta}]\!]$. (Hint: try $\alpha + \beta = 0$.)

7.17 Show that the map $z \mapsto \theta(z)$ (where $0 \neq x = |z|\mathrm{e}^{\mathrm{i}\theta(z)}$) is continuous in $\mathbb{C}_\pi = \mathbb{C} \setminus [0,\infty)$ if $\theta(z)$ is chosen to lie in $(0, 2\pi)$. (Hint: consider a disc $\mathrm{D}(a;r)$ within $\mathbb{C}_\pi$ and argue geometrically or use the fact that $\tan\theta(z) = y/x$, where $z = x + \mathrm{i}y$.)

# 8 Conformal mapping

This chapter concerns angle-preserving mappings between regions in the complex plane. As we shall see, every holomorphic function whose derivative is non-zero defines such a mapping. These mappings are of intrinsic geometric interest and of importance in advanced complex analysis (Chapter 16 hints at this). They are also worth studying because of their usefulness in solving certain physical problems, for example, problems about two-dimensional fluid flow, the idea being to transform a given problem into an equivalent one which is easier to solve. So we wish to consider the problem of mapping a given region $G$ onto a geometrically simpler region $G'$, for example the open unit disc or the open upper half-plane. We concentrate in this basic track chapter on presenting and illustrating the principles of conformal mapping. We therefore restrict attention to mapping a region whose boundary is a circline or is a pair of arcs (that is, lines, rays, circular arcs, line segments). For many such problems, three basic types of map (and certain combinations of these) suffice:

- Möbius transformations (studied quite extensively in Chapter 2),
- integer and non-integer powers, and
- exponentials (and their inverses, logarithms).

We must take due care to avoid many-valuedness whenever we consider logarithms or non-integer powers. Otherwise in this chapter we can safely treat argument 'mod $2\pi$' and so let $\arg z$ denote any choice from the set $[\![ \arg z ]\!]$.

We have omitted trigonometric and hyperbolic functions from our catalogue of maps because the uses of such maps are relatively specialized. We consider them in Chapter 23; this is geared to applications and includes a variety of further examples of mappings.

## Conformal mapping

Here we reveal the connection between angle-preservation and holomorphy.

**8.1 Angles between paths.** Let $\gamma$ be a path with, for definiteness, parameter interval $[0,1]$. Then there is a well-defined tangent to $\gamma$ at $\zeta = \gamma(0)$. This is defined by $\zeta + t\gamma'(0)$ $(t \geqslant 0)$, provided $\gamma'(0) \neq 0$, and makes an angle $\arg \gamma'(0)$ with the real axis.

Let $\gamma_1$ and $\gamma_2$ be paths, both with parameter interval $[0, 1]$, having common endpoint $\zeta = \gamma_1(0) = \gamma_2(0)$. We assume that $\gamma_1'(0)$ and $\gamma_2'(0)$ are non-zero, so that $\gamma_1$ and $\gamma_2$ have well-defined tangents at $\zeta$, with the angle between $\gamma_1$ and $\gamma_2$ being (by definition) $\arg \gamma_1'(0) - \arg \gamma_2'(0)$.

**8.2 Conformality theorem.** Suppose that $f$ is holomorphic in an open set $G$, that $\gamma_1$ and $\gamma_2$ are paths (with parameter interval $[0, 1]$) in $G$ meeting at $\zeta = \gamma_1(0) = \gamma_2(0)$, and that $f'(\zeta) \neq 0$. Then, in the sense defined below, $f$ preserves angles between paths in $G$ meeting at $\zeta$.

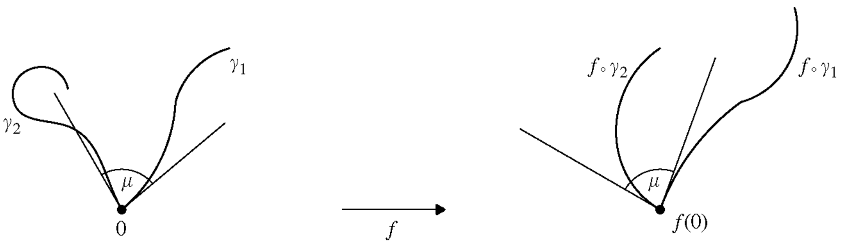



**Figure 8.1** Preservation of an angle under a conformal map

**Proof** Let the angle between $\gamma_1$ and $\gamma_2$ be $\lambda = \arg \gamma_1'(0) - \arg \gamma_2'(0)$. The paths $\gamma_1$ and $\gamma_2$ are mapped by $f$ to paths $f \circ \gamma_1$ and $f \circ \gamma_2$; note that these are indeed paths. They meet at $f(\zeta)$ at an angle $\Lambda = \arg(f \circ \gamma_1)'(0) - \arg(f \circ \gamma_2)'(0)$. The assertion of the theorem is that $\Lambda = \lambda \pmod{2\pi}$. By the Chain rule,

$$\frac{(f \circ \gamma_1)'(0)}{(f \circ \gamma_2)'(0)} = \frac{f'(\zeta)\gamma_1'(0)}{f'(\zeta)\gamma_2'(0)} = \frac{\gamma_1'(0)}{\gamma_2'(0)},$$

from which the result follows; see 7.13 concerning args of quotients. □

**8.3 Conformal mapping.** A complex-valued function is **conformal in an open set** $G \subseteq \mathbb{C}$ (or $\widetilde{\mathbb{C}}$) if $f \in \mathrm{H}(G)$ and $f'(z) \neq 0$ for every $z \in G$. It is **conformal at a point** $\zeta$ if it is conformal in some $\mathrm{D}(\zeta; r)$.

The Conformality theorem shows that a conformal mapping preserves both the magnitude and sense of angles between paths; informally, preservation of

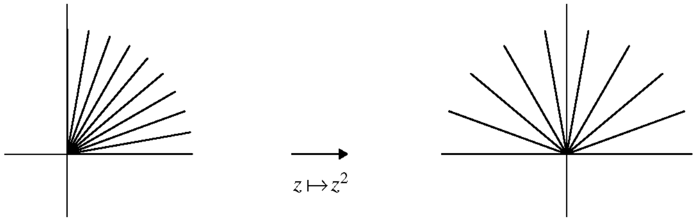



**Figure 8.2** Non-conformality of $z \mapsto z^2$ at 0

sense means that orientation is preserved: see Fig. 8.1. For a converse to the Conformality theorem, see Exercise 8.15.

We should not be surprised at the restriction $f'(\zeta) \neq 0$ in the Conformality theorem. The function $z \mapsto z^2 = w$ takes rays $\arg z = \lambda$ and $\arg z = \mu$ meeting at angle $\lambda - \mu$ to rays $\arg w = 2\lambda$ and $\arg w = 2\mu$ meeting at angle $2(\lambda - \mu)$. See Fig. 8.2 and also Exercise 8.3.

**8.4 Construction of conformal mappings: preliminary remarks.** Suppose we require a conformal map $f$ from the open upper half-plane $\Pi^+ = \{ z : \operatorname{Im} z > 0 \}$ onto the open unit disc $\mathrm{D}(0;1)$. It is unlikely to be helpful to bring $\operatorname{Im} z$ into the definition of $f$, since $\operatorname{Im} z$ is not holomorphic and we want $f \in \mathrm{H}(\Pi)^+$. But compare the following descriptions:

$$\Pi^+ = \{ z : |z - \mathrm{i}| < |z + \mathrm{i}| \} = \left\{ z : \left| \frac{z - \mathrm{i}}{z + \mathrm{i}} \right| < 1 \right\},$$
$$\mathrm{D}(0;1) = \{ w : |w| < 1 \}.$$

It ought now to be entirely obvious that the map we want is

$$f\colon z \mapsto w = (z - \mathrm{i})/(z + \mathrm{i})$$

since

$$z \in \Pi^+ \iff f(z) \in \mathrm{D}(0;1).$$

Also, $f$ is holomorphic in $\Pi^+$ with derivative $f'(z) = 2\mathrm{i}(z + \mathrm{i})^{-2} \neq 0$, and so is conformal. Notice that the non-holomorphic function $|z|$ enters into the descriptions we use to match up the domain and image regions, but does not enter into the definition of the function $f$. This simple example shows that success in constructing a conformal map from one region to another will depend on a judicious choice of descriptions for the regions. Note also that use of the

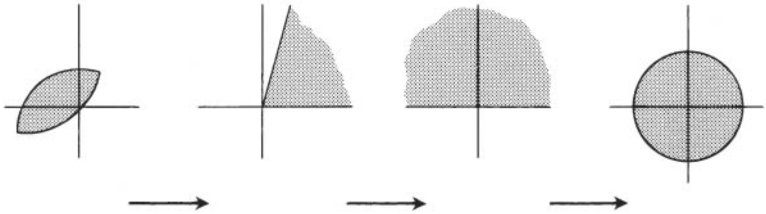


**Figure 8.3** A multi-stage map

'matching up' technique guarantees that the constructed map takes the domain *onto*, and not merely into, the target region.

The Chain rule implies that the composition of conformal maps is itself conformal. This elementary observation is very useful: it allows us to build up conformal mappings in a finite number of simple steps. For example, to map a lozenge onto $D(0;1)$ we might proceed as indicated in Fig. 8.3. Thus an aid to successful map-building is familiarity with standard mappings.

A plan for mapping a given region $G$ onto, say, $D(0;1)$ might go as follows. Try to map $G$ first to some more familiar region $G'$, in the knowledge that, having reached $G'$, a sequence of 'off-the-peg' maps will then take us to $D(0;1)$. Giving a distinguished role to $D(0;1)$ is justified by the following strategy for mapping one region $G_1$ onto another, $G_2$:

$$G_1 \xrightarrow{\ f_1\ } D(0;1) \underset{f_2^{-1}}{\overset{f_2}{\rightleftarrows}} G_2.$$

If we can map $G_1$ onto $D(0;1)$ using a conformal map $f_1$ and $G_2$ onto $D(0;1)$ using a map $f_2$ for which the inverse exists and is conformal, then $f_2^{-1} \circ f_1$ maps $G_1$ conformally onto $G_2$. Theorems guaranteeing holomorphy and conformality of inverse maps appear in Chapter 16.

## Some standard conformal mappings

In this section we study the mapping properties of Möbius transformations, exponentials, and powers.

We have already investigated Möbius transformations quite thoroughly in Chapter 2; you are advised to review this material before proceeding. The good thing about Möbius maps is that they map circlines to circlines and arcs to arcs, and so are a natural choice for mapping regions bounded by circlines and arcs. We saw in Chapter 2 that Möbius transformations are best viewed as mappings of the extended plane, $\widetilde{\mathbb{C}}$. We continue to work in $\widetilde{\mathbb{C}}$ when appropriate.

**8.5 Particularly useful Möbius mappings.** All half-planes below are open half-planes.

- Upper half-plane onto unit disc: $z \mapsto \dfrac{z - \mathrm{i}}{z + \mathrm{i}}$.
- Lower half-plane onto unit disc: $z \mapsto \dfrac{z + \mathrm{i}}{z - \mathrm{i}}$.
- Right half-plane onto unit disc: $z \mapsto \dfrac{z - 1}{z + 1}$ (and vice versa; this map is self-inverse).
- Left half-plane onto unit disc: $z \mapsto \dfrac{z + 1}{z - 1}$.

In addition, the boundary line of each half-plane maps onto the unit circle.

Since we use Möbius mappings so often, it is worth knowing that they are conformal everywhere.

**8.6 Conformality of Möbius transformations.** Consider the Möbius transformation $f\colon z \mapsto (az + b)/(cz + d)$ $(ad - bc \neq 0)$. Certainly, provided that $cz + d \neq 0$,

$$f'(z) = \frac{ad - bc}{(cz + d)^2} \neq 0.$$

Hence $f$ is conformal in $\mathbb{C} \setminus \{-d/c\}$ $(c \neq 0)$.

We normally regard a Möbius transformation as a mapping from $\widetilde{\mathbb{C}}$ to itself, so want to extend our conformality definition to such a mapping. If $f$ maps a point $\zeta \in \mathbb{C}$ to $\infty$, then we consider $g\colon z \mapsto 1/f(z)$ and say that $f$ is conformal at $\zeta$ if $g$ is conformal at $\zeta$. We also say that $f$ is conformal at $\infty$ if $\widetilde{f}$, defined as usual by $\widetilde{f}(z) = f(1/z)$, is conformal at $z = 0$. Take $c \neq 0$ and consider $\widetilde{f}$:

$$\widetilde{f}(\zeta) = \frac{b\zeta + a}{d\zeta + c}.$$

From above, this is conformal at 0, so $f$ is conformal at $\infty$. Now consider the behaviour at $z = -d/c$, which $f$ maps to $\infty$. To see what happens, let $\tau = 1/w$ where $w = f(z)$. Then $\tau = (cz+d)/(az+b)$, and this has a non-zero derivative at $z = 0$. Finally, in the case when $c = 0$, we have $f(\infty) = \infty$. By considering $\tau = 1/w$ as a function of $\zeta = 1/z$, it is easy to show that the derivative at $z = 0$ is non-zero. We conclude that a Möbius transformation has a non-zero derivative at every point of $\widetilde{\mathbb{C}}$.

We can view geometrically the conformality of a Möbius transformation $f$ with $f(-d/c) = \infty$, where $c \neq 0$. For example, $f$ will map a pair of circles having a common tangent at $-d/c$ to a pair of parallel lines.

**8.7 Mappings by a positive integer power.** The map $z \mapsto z^n$ $(n = 2, 3, \dots)$ is conformal except at 0, where angles between paths are magnified by a factor of $n$. The non-conformality at 0 can be an asset rather than a snag, so long as 0 lies outside the region being mapped. For example, $z \mapsto z^2$ maps

- a quadrant to a half-plane and, in particular, the first quadrant

$$\{ z : \operatorname{Re} z > 0, \operatorname{Im} z > 0 \}$$

  onto the open upper half-plane $\Pi^+$, and

- $\Pi^+ = \{ z : 0 < \arg z < \pi \}$ onto $\{ z : 0 < \arg z < 2\pi \} = \mathbb{C} \setminus [0, \infty)$.

Unlike a Möbius transformation, $z \mapsto w = z^2$ does not map arcs or circlines to curves of the same type. It is easy to verify that, in general, $z \mapsto z^2$ maps a line to a parabola.

**8.8 Mapping by a general power.** Here we have to contend with a multifunction. Take $\alpha > 0$ and consider, for definiteness, $z^\alpha = |z|^\alpha \, e^{i\alpha\theta}$ $(z = |z| \, e^{i\theta}$, $0 < \theta < 2\pi)$. Then $z \mapsto z^\alpha$ is conformal in the plane cut along $[0, \infty)$. See Fig. 8.4 for an illustration.

The position of the cut may be governed by the region we wish to consider. For example, if we wished to find a conformal mapping from $\mathbb{C} \setminus (-\infty, 0]$ onto the open right half-plane, we would cut the plane along $(-\infty, 0]$ and take $z \mapsto z^{1/2} = |z|^{1/2} \, e^{i\theta/2}$, where $z = e^{i\theta}$ $(-\pi < \theta < \pi)$.

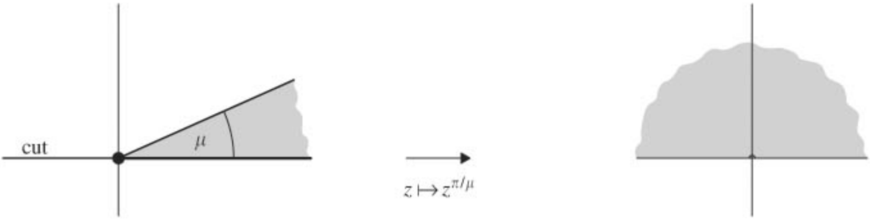



**Figure 8.4** Mapping of a sector onto $\Pi^+$

**8.9 Mapping by exponentials and logarithms.** Let

$$f\colon z = x + \mathrm{i}y \mapsto \mathrm{e}^z = w = R\mathrm{e}^{\mathrm{i}\varphi}.$$

This map is conformal everywhere, by 7.3. Then (recall 7.4)

$$R = \mathrm{e}^x \quad \text{and} \quad \varphi = y \pmod{2\pi}.$$

Hence $z \mapsto \mathrm{e}^z$ maps

| | | |
|---|---|---|
| the vertical line $x = a$ | to | the circle $\lvert w\rvert = \mathrm{e}^a$, |
| the horizontal line $y = c$ | to | the ray $\arg w = c$. |

Therefore $z \mapsto \mathrm{e}^z$ maps

| | | |
|---|---|---|
| the vertical strip $\{\, z : a < \operatorname{Re} z < b \,\}$ | to | the annulus $\{\, w : \mathrm{e}^a < \lvert w\rvert < \mathrm{e}^b \,\}$, |
| the horizontal strip $\{\, z : c < \operatorname{Im} z < d \,\}$ | to | the sector $\{\, c < \arg w < d \,\}$. |

See Fig. 8.5.

In reverse, a logarithm will map a sector to a strip, but we must work in a cut plane and select a holomorphic branch. The cut must be placed so that it does not encroach on the region we wish to map.

## Mappings of regions by standard mappings

Although it is important to know how various maps treat curves in $\mathbb{C}$ or in $\widetilde{\mathbb{C}}$, in many applications we need to know about mappings of regions, with or without their boundaries, and we shall increasingly study regions rather than their boundary curves.

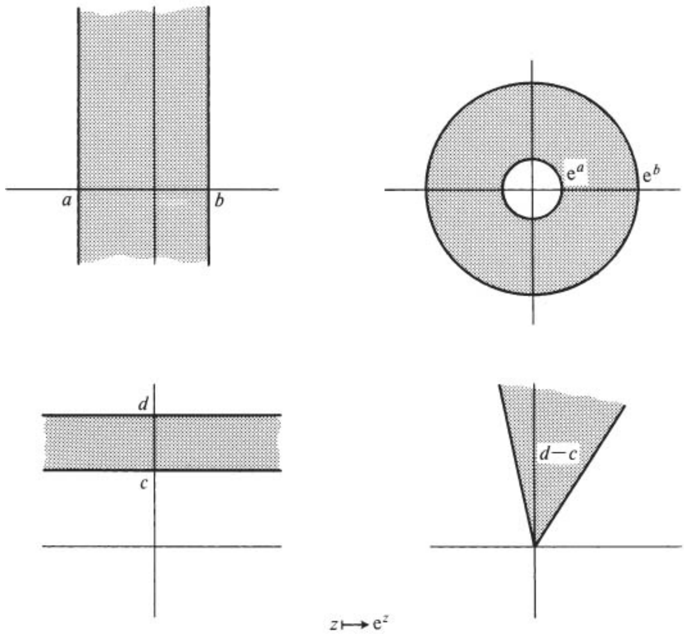



**Figure 8.5** Mappings of strips by $z \mapsto e^z$

**8.10 Mappings of regions bounded by arcs.** Consider a lozenge-shaped region $G$ as shown in Fig. 8.6. Take the Möbius transformation

$$f: z \mapsto \frac{z-\alpha}{z-\beta}.$$

This sends $\alpha$ to 0 and $\beta$ to $\infty$ and maps any arc joining $\alpha$ and $\beta$ to a ray from 0 to $\infty$. The boundary arcs meet at $\alpha$ at an angle $\mu$ so, by conformality of $f$, the boundary rays of the image meet at 0 at the same angle. Therefore the image is a sector of angle $\mu$. The image of a single point other than $\alpha$ or $\beta$ uniquely determines the position of the image sector. A rotation $w \mapsto e^{i\varphi}w$ will swing the sector round to any desired position. Thus by taking the map $z \mapsto k(z-\alpha)/(z-\beta)$, where $k$ is a constant of modulus 1, we can map $G$ conformally onto any desired sector of angle $\mu$.

It is entirely plausible that the region between the bounding arcs of $G$ is mapped to the region bounded by their images, as we tacitly assumed above. [Fully validating this claim involves topological arguments using the fact that the image of a connected set under a continuous map is connected.] We can avoid having to decide whether a particular conformal map takes a region interior to a closed path to the inside or to the outside of the image of the path by working from the start with the given region rather than with just the boundary.

In the present example we can view $G$ and its boundary as a union of circular arcs joining $\alpha$ to $\beta$. Each such arc maps to a ray starting at 0, and the image of $G$ is an open sector formed by a fan of rays. This is bounded by the images of the boundary arcs of $G$ which, as we have already said, meet at an angle $\mu$.

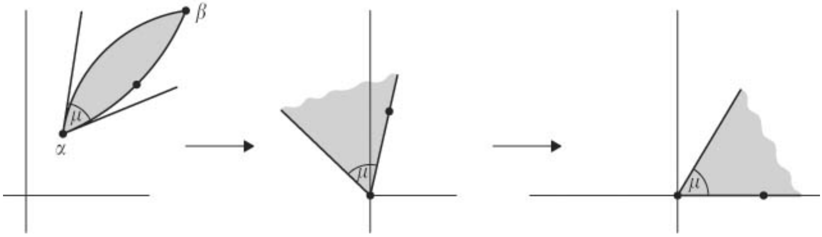



**Figure 8.6** Mapping a lozenge to a sector

**8.11 Example (mapping a semicircular region onto a disc).** Here is a concrete example illustrating the technique in 8.10. Let $f\colon z \to w = (z+1)/(z-1)$ and consider the image under $f$ of the semicircular region

$$S := \{\, z \in \mathbb{C} : |z| < 1,\ \operatorname{Im} z > 0 \,\};$$

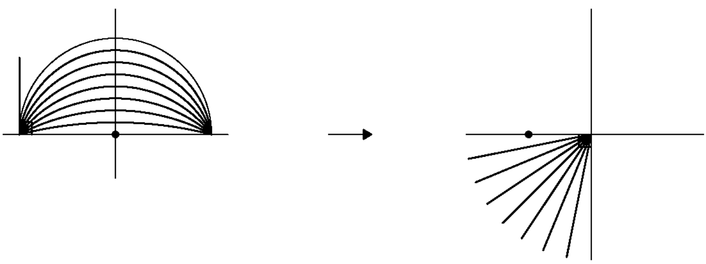


**Figure 8.7** Example 8.11

The map $f$ sends $-1$ to $0$ and $0$ to $\infty$, and so maps arcs through $-1$ and $1$ to rays from $0$ to $\infty$. Denote $-1$, $1$, $0$, $\mathrm{i}$ by $A$, $B$, $O$, $P$. The boundary arc $APB$ and segment $AOB$ of $S$ meet at an angle $\pi/2$. Since $f$ is conformal at $-1$, the angle between the images is $\pi/2$ as well. Also, the point $-1 = f(0)$ lies on the image of $AOB$, so this image ray must be the negative real axis. Because $f$ has to preserve the sense of angles, the image of $APB$ must be the non-positive imaginary axis.

We can confirm by techniques from Chapter 2 that the above reasoning has led to the correct conclusion. An alternative description of $S$ is

$$S := \{ z \in \mathbb{C} : \arg((z+1)/(z-1)) = \mu, \text{ where } -\pi < \mu < -\pi/2 \}$$

(cf. 2.5). It is then immediate that

$$f(S) = \{ w \in \mathbb{C} : -\pi < \arg w < -\pi/2 \}.$$

## Tactical tips

- ⊙ As we have just illustrated, it is possible, using the results in 2.4, to find an *explicit* equation for any given arc through points $\alpha$ and $\beta$ in terms of the angle subtended. However, such explicit representations can be tiresome to find, and great care must be taken with signs. Hence it is usually preferable to employ geometric arguments rather than analytic ones based on arg equations for arcs.
- ⊙ At this stage, finding images by the substitution method may seem easier and more reliable than exploiting conformality. However, working with arg equations of arcs or with inverse-point representations of circlines can be tedious. Also, we shall shortly want to *construct* conformal maps between specified regions, rather than finding images under *given* maps. For this, substitution is not an option and geometric thinking is highly recommended.

**8.12 Example (mapping a region bounded by non-concentric circles).** Suppose we have a region $G$ as shown in Fig. 8.8. It is an easy exercise to show that any pair of non-concentric circles with one lying strictly inside the other has a common pair of inverse points, $\alpha$ and $\beta$. (That is, the given circles form two members of a coaxal system $\mathcal{C}_1(\alpha, \beta)$; see 2.12.) The equations of the two circles in inverse-point form are

$$\left| \frac{z-\alpha}{z-\beta} \right| = \lambda_1 \quad \text{and} \quad \left| \frac{z-\alpha}{z-\beta} \right| = \lambda_2,$$

where we assume $\lambda_1 < \lambda_2$. The map $z \mapsto w := (z-\alpha)/(z-\beta)$ takes the region $G$ onto the annulus $\{ w : \lambda_1 < |w| < \lambda_2 \}$.

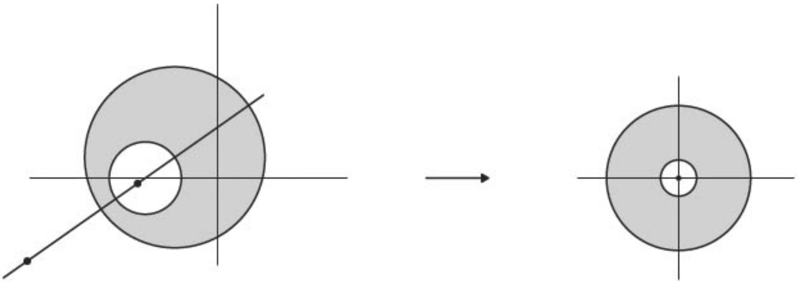


**Figure 8.8** Mapping to an annulus

**8.13 Example (mapping a region bounded by touching circles).** Consider the region $G$ bounded by the circles $C_0$ and $C_1$ given by $|z-1| = 1$ and $|z| = 2$. We find the image of $G$ under the Möbius transformation $f\colon z \mapsto w = 1/(z-2)$.

We can view $G$ as the union of circles $C_a$ with equation $|z-1+a| = 1+a$ $(0 < a < 1)$. Write $z = (1+2w)/w$. The substitution method gives the equation of $f(C_a)$ to be $|w + 1/(1+a)| = |w|$. Thus the image of $C_a$ is a vertical line $\operatorname{Re} z = -1/2(1+a)$ and the image region is the vertical strip $G'$ given by $-1/2 < \operatorname{Re} w < -1/4$.

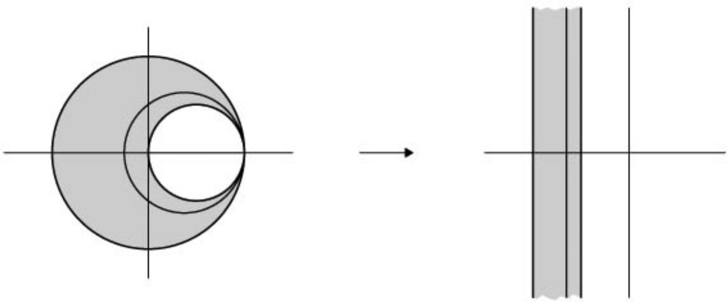


**Figure 8.9** Example 8.13

Alternatively, we may argue geometrically. The boundary circles $C_0$ and $C_1$ touch at $z = 2$, where the angle between their tangents is 0. The images of $C_0$ and $C_1$ are circlines (by Theorem 2.15). They pass through $f(2) = \infty$ and so are straight lines. Because $f$ is conformal, the angle between these image lines is 0, that is, the lines are parallel. The real axis cuts $C_0$ and $C_1$ orthogonally. Hence its image (easily seen to be the real axis) cuts $f(C_0)$ and $f(C_1)$ orthogonally, by conformality of $f$. These lines are therefore vertical and pass through $f(0) = -1/2$ and $f(-2) = -1/4$. The point $-1$ lies inside $G$ and is mapped to $-1/3 \in G'$. We conclude that the image of $G$ is indeed $G'$.

## Building conformal mappings

So far, we have given examples of the effect of a given map on a region. In practice, we often need to be able to construct a conformal map of some given region onto a simpler one such as a disc or a half-plane.

**8.14 Example (mapping a quadrant of a disc onto a disc).** We find a conformal mapping of $G = \{ z : |z| < 1,\ \operatorname{Im} z > 0, \operatorname{Re} z > 0 \}$ onto $D(0; 1)$.

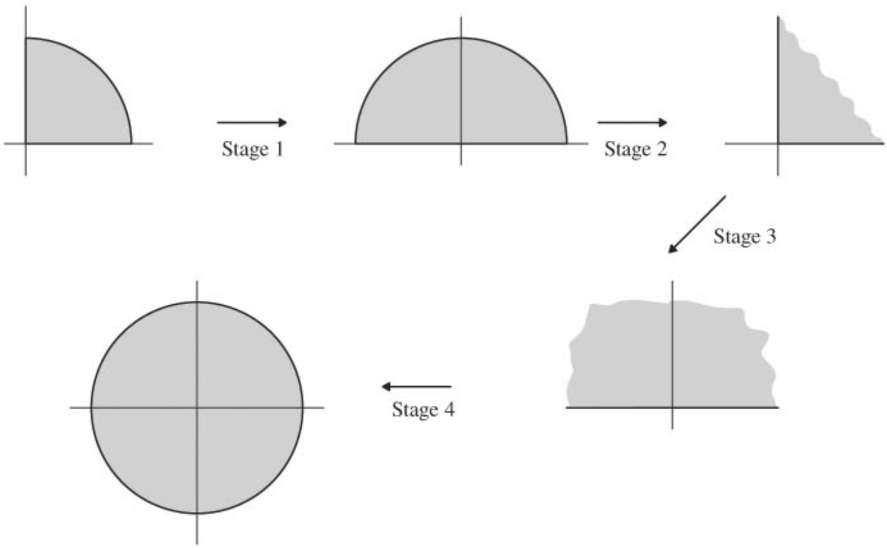



**Figure 8.10** From the quadrant of a disc to $D(0; 1)$

**Stage 1** Note that $G$ consists of those points $z$ with $|z| < 1$ and $0 < \arg z < \pi/2$. Also $|z| < 1$ if and only if $|z|^2 < 1$. Hence $z \mapsto w = z^2$ maps $G$ conformally onto the semicircular region $S$ considered in Example 8.11.

**Stage 2** A map of the form

$$w \mapsto \zeta = k\frac{w+1}{w-1} \quad (|k| = 1)$$

will map $S$ conformally onto a quadrant $Q$. By choosing $k = -1$, we send $0$ to $1$, and this makes $Q$ the first quadrant.

**Stage 3** Let $\zeta \mapsto \tau = \zeta^2$. This maps the quadrant $Q$ conformally onto $\Pi^+$.

**Stage 4** Let $\tau \mapsto \eta = (\tau - \mathrm{i})/(\tau + \mathrm{i})$ to map $\Pi^+$ conformally onto $\mathrm{D}(0;1)$.

Putting these maps together, we let $f\colon z \mapsto \eta$, where

$$\eta = \frac{((z^2+1)/(1-z^2))^2 - \mathrm{i}}{((z^2+1)/(1-z^2))^2 + \mathrm{i}} = \frac{(z^2+1)^2 - \mathrm{i}(1-z^2)^2}{(z^2+1)^2 + \mathrm{i}(1-z^2)^2} = \mathrm{i}\frac{z^4 + 2\mathrm{i}z^2 + 1}{z^4 - 2\mathrm{i}z^2 + 1}.$$

By construction, $f$ maps $G$ onto $\mathrm{D}(0;1)$. As the composite of conformal maps, $f$ is conformal.

### Tactical tip

- ⊙ Could we simply have used a power $f\colon z \mapsto z^n$ to map $G$ onto $\mathrm{D}(0;1)$? For $n \geqslant 5$, this nearly succeeds: $f(S) = \mathrm{D}'(0;1)$. However, since $0 \notin G$, no power will map $G$ onto $\mathrm{D}(0;1)$.

**8.15 Example (mapping a crescent onto $\mathrm{D}(0;1)$).** Consider again the region $G$ bounded by the circles $|z| = 2$ and $|z-1| = 1$.

**Stage 1** Our first move is to map the 'awkward' point $z = 2$ to $\infty$ by taking the map $z \mapsto w = 1/(z-2)$. The image of $G$ in the $w$-plane is then the strip $G'$ given by $-1/2 < \operatorname{Re} w < -1/4$.

**Stage 2** We now transform $G'$ to a strip more amenable to mapping by an exponential. We let $\zeta = 4\pi\mathrm{i}(w + \frac{1}{2})$. Then the image of $G'$ in the $\zeta$-plane is given by $0 < \operatorname{Im}\zeta < \pi$.

**Stage 3** We can transform the strip from Stage 2 to the open upper half-plane by means of $\zeta \mapsto \tau = \mathrm{e}^{\zeta}$ (see 8.9).

**Stage 4** Finally, we map the open upper half-plane to $\mathrm{D}(0;1)$ using $\tau \mapsto \eta = (\tau - \mathrm{i})/(\tau + \mathrm{i})$ (see 8.5).

Therefore $z \mapsto \eta$ maps $G$ onto $\mathrm{D}(0;1)$ and is conformal, since each of the maps in Stages 1–4 is conformal on its domain. We can, if required, compute $\eta$ as a function of $z$.

## Exercises

**Exercises from the text.** Verify that a pair of circles, one of which lies inside the other, has a common pair of inverse points (8.10).

8.1 Find the image of

$$\text{(i) } \{\, z : 0 < \arg z < \pi/4 \,\}, \quad \text{(ii) } D'(0;2), \quad \text{(iii) } \{\, z : 0 < \operatorname{Im} z < 1 \,\}$$

under (a) $z \mapsto (1+\mathrm{i})z$, (b) $z \mapsto 1/z$. (Argue geometrically whenever you can.)

8.2 Find the image of
(i) $\{\, z : 0 < \arg z < 2\pi/3 \,\}$ under $z \mapsto z^5$,
(ii) $\{\, z : \operatorname{Im} z > 0, \operatorname{Re} z > 0 \,\}$ under $z \mapsto z^3$.

8.3 Consider the map $z \mapsto z^2$.
(i) Find the image of $\{\, z : 0 < \operatorname{Im} z < 1 \,\}$.
(ii) Find, and sketch, the images of the lines $\operatorname{Im} z = \mu$ for $0 < \mu < 1$.

8.4 Find the image of
(i) $\{\, z : -\pi < \operatorname{Im} z < \pi/2 \,\}$ under $z \mapsto \mathrm{e}^z$,
(ii) $\{\, z : -1 < \operatorname{Re} z < 1 \,\}$ under $z \mapsto \mathrm{e}^{\mathrm{i}\pi z}$,
(iii) $\{\, z : \operatorname{Re} z > 0 \,\}$ under $z \mapsto \mathrm{e}^z$.

8.5 (A refresher on Möbius transformations, for which a variety of techniques is recommended.) Describe the image of
(i) $\{\, z : |z-1| > 1 \,\}$ under $z \mapsto w = z/(z-2)$,
(ii) $\{\, z : \frac{1}{2} < |z| < 1 \,\}$ under $z \mapsto w = (2z+1)/(z-2)$,
(iii) $\{\, z : \operatorname{Re} z > 0 \,\}$ under $z \mapsto w$, where $(w-1)/(w+1) = 2(z-1)/(z+1)$,
(iv) $\{\, z : |z-\mathrm{i}| < 1, \operatorname{Re} z < 0 \,\}$ under $z \mapsto (z-2\mathrm{i})/z$,
(v) $D(0;1)$ under $z \mapsto (z-\frac{1}{2})/(z-2)$.

8.6 Find a common pair of inverse points for the circles $|z| = 1$ and $|z-1| = \frac{5}{2}$. Hence find a conformal mapping of the region bounded by these two circles onto an annulus.

8.7 Describe the image of
(i) $\{\, z : |z-1| > 1, |z+1| > 1 \,\}$ under $z \mapsto w = \mathrm{i}(z+2)/z$,
(ii) $\{\, z : 0 < \arg z < \pi/4 \,\}$ under $z \mapsto w = z/(z+1)$,
(iii) $\{\, z : 0 < \arg z < \pi/4 \,\}$ under $z \mapsto w = z/(z-1)$ (hint: this is harder than (ii) and it may help to recall Exercise 2.6),

(a) by the substitution method and (b) by arguing geometrically.

8.8 Find the image of $\{ z : \operatorname{Re} > 0,\ \operatorname{Im} z > 0,\ |z| > 1 \}$ under $z \mapsto \log z = \log|z| + i\theta$, where $z = |z|e^{i\theta}$ $(-\pi < \theta \leqslant \pi)$ and the plane is cut along $(-\infty, 0]$.

8.9 By considering the map as a composite of simpler maps, find the image of

(i) $\{ z : 0 < \arg z < \pi/4 \}$ under $z \mapsto w = iz^4$,

(ii) $\{ z : 0 < \operatorname{Im} z < \pi \}$ under $z \mapsto w = \dfrac{1 + ie^z}{1 - ie^z}$,

(iii) $\{ z : |z| < 1,\ 0 < \arg z < \pi/3 \}$ under $z \mapsto w = \left(\dfrac{z^3+1}{z^3-1}\right)^2$,

(iv) $\{ z : |z| < 1,\ \left|z - \frac{1}{2}\right| > \frac{1}{2} \}$ under $z \mapsto w = e^{1/(z-1)}$,

(v) $\{ z : |z - i| < \sqrt{2},\ |z + i| < \sqrt{2} \}$ under $z \mapsto \left(\dfrac{z+1}{z-1}\right)^2$.

8.10 Construct a conformal mapping onto $D(0;1)$ of each of the following:

(i) $\{ z : -1 < \operatorname{Re} z < 1 \}$, (ii) $\{ z : 0 < \operatorname{Im} z < 2 \}$,

(iii) $\{ z : -\frac{1}{4}\pi < \arg z < \frac{3}{4}\pi \}$, (iv) $\{ z : \operatorname{Re} z > 0 \text{ or } \operatorname{Im} z \neq 0 \}$.

8.11 Construct a conformal mapping of $G_1$ onto $G_2$ when

(i) $G_1 = \{ z : \operatorname{Im} z < \frac{1}{2} \}$ and $G_2 = D(1;1)$,

(ii) $G_1 = \{ z : -\frac{\pi}{2} < \arg z < \frac{\pi}{2} \}$ and $G_2 = \{ w : |w| < 1,\ \operatorname{Im} w < 0 \}$,

(iii) $G_1 = \{ |\operatorname{Im} z| < \frac{\pi}{6} \}$ and $G_2 = \{ w : \operatorname{Re} w > 0 \}$.

8.12 Given that $-1 < c < 1$, find a conformal mapping of

$$\{ z : |z| < 1,\ \operatorname{Re} z > c \}$$

onto the open upper half-plane.

8.13 Let $U := \mathbb{C} \setminus \{ z : |z| = 1,\ \operatorname{Im} z \geqslant 0 \}$.

(i) Find the image of $\mathbb{C} \setminus [-1, 1]$ under $z \mapsto (z-1)/(z+1)$.

(ii) Hence find a conformal mapping of $U$ onto a half-plane.

8.14 Show that $z \mapsto \zeta = i(1+z)/(1-z)$ maps the unit circle onto the real axis and deduce that the same is true of the map $z \mapsto \zeta^3$.
Determine, and sketch, the regions in the $z$-plane that are carried onto the right half-plane $\{ w : \operatorname{Re} w > 0 \}$ by (i) $z \mapsto \zeta$, (ii) $z \mapsto \zeta^3$.

8.15 Let $G$ be an open set and let $\gamma$ be a path with $\gamma^* \subseteq G$. Suppose that $f\colon G \to \mathbb{C}$ is a function such that the partial derivatives $f_x$ and $f_y$ exist and are continuous in $G$ and let $\Gamma = f \circ \gamma$ be the image of $\gamma$ under $f$. Show that

$$\Gamma'(t) = \frac{1}{2}(f_x - \mathrm{i} f_y)\gamma'(t) + \frac{1}{2}(f_x + \mathrm{i} f_y)\overline{\gamma'(t)},$$

where the partial derivatives are evaluated at $\gamma(t)$. By considering

$$\arg\big(\Gamma'(t)/\gamma'(t)\big)$$

for suitable choices of $\gamma$, show that, if $f$ preserves the magnitude and sense of angles between paths in $G$, then the real and imaginary parts of $f$ satisfy the Cauchy–Riemann equations. (Hence, by 5.6, $f \in \mathrm{H}(G)$.)

# 9 Multifunctions

Multifunctions are hard to avoid. Many complex functions, like the complex exponential, are not globally one-to-one. We may view such a function as having an inverse, so long as we allow the inverse to be a multifunction. Constructing, at least locally, a well-behaved functional inverse will involve extracting a suitable value from this multifunction at each point of the domain. We have already seen how to make a selection of values from the logarithm and from a fractional power of $z$ to arrive at a holomorphic function. But how should we treat more complicated examples, such as $[\![\log\big((1+z)/(1-z)\big)]\!]$ or $[\![(z^3-1)^{1/2}]\!]$? We now begin a deeper analysis of many-valuedness, which will enable us to handle logarithms and powers of rational functions. Multifunctions of this sort reappear in Chapter 23, in connection with conformal mappings arising in applied mathematics.

This chapter, on the basic track, explains how to handle multifunctions in a way which is mathematically sound but, deliberately, unsophisticated. A more formal and more elegant treatment of argument can be found in Chapter 12.

## Branch points and multibranches

**9.1 Argument as a function.** The restriction $-\pi < \theta \leqslant \pi$ or, alternatively, $0 \leqslant \theta < 2\pi$, uniquely determines $\theta$ in the equation $0 \neq z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$.

Now consider what happens to a principal-value determination of argument, $\operatorname{Arg} z = \theta$, where $z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$, $0 \leqslant \theta < 2\pi$, when $z$ performs a complete anticlockwise circuit round the unit circle starting from $z = 1$, with $\theta \in [0, 2\pi)$. Within the chosen range $[0, 2\pi)$, $\theta$ has value $0$ at the start and increases steadily towards $2\pi$ as $z$ moves round the circle until it arrives back at $1$, when $\theta$ must jump back to $0$. Thus $\operatorname{Arg} z$ has a jump discontinuity. On the other hand, if we insist on choosing $\theta$ so that it varies continuously with $z$, then its final value has to be $2\pi$, a different choice from $[\![\arg 1]\!]$ from that we made at the start.

We can give a more formal treatment of the issues just discussed.

**9.2 The lack of a continuous argument function.** We show that there is *no* way to impose a restriction which selects $\theta(z) \in [\![\arg z]\!]$, for *all* $z$ in $\mathbb{C} \setminus \{0\}$, so that $\theta \colon z \mapsto \theta(z)$ is continuous as a function of $z$. We assume for a contradiction that such a continuous function $\theta$ does exist and consider

$$k(t) = \frac{1}{2\pi}\big(\theta(e^{it}) + \theta(e^{-it})\big) \quad (t \in \mathbb{R}).$$

Then $k$ is continuous and

$$k(t) = \frac{1}{2\pi}\big((t + 2m_t\pi) + (-t + 2n_t\pi)\big) \qquad \text{where } m_t, n_t \in \mathbb{Z},$$

so $k$ takes only integer values. Also $k(0)$ is even and $k(\pi)$ is odd, so $k$ is non-constant. This contradicts Theorem 3.25 (a consequence of the real Intermediate value theorem).

This result has implications for other multifunctions. For example, it tells us that there cannot be a continuous logarithm in $\mathbb{C} \setminus \{0\}$: if there were one, then its imaginary part—an argument function—would be continuous too.

**9.3 Branch points.** Take a multifunction $[\![w(z)]\!]$ (so that $w(z)$ is a non-empty subset of $\mathbb{C}$ for each $z$ in the domain of definition of $w$). Assume that the many-valuedness arises because, for one or more points $a$, the definition of $w(z)$ explicitly or implicitly involves the angle $\theta$, where $z - a = |z - a|\, e^{i\theta}$. Such points are called **branch points**. Any branch point is excluded from the domain of definition of $[\![w(z)]\!]$. More formally, $a$ is a branch point for $[\![w(z)]\!]$ if, for all sufficiently small $r > 0$, it is not possible to choose $f(z) \in [\![w(z)]\!]$ so that $f$ is a continuous function on $\gamma(a; r)^*$. The motivation comes from 9.2: no continuous argument function can be drawn from $[\![\arg(z - a)]\!]$ for $z$ on a circle centre $a$.

**9.4 Examples (branch points).**

- $[\![\log z]\!]$ has a branch point at 0.
- $[\![(z-1)^{1/2}]\!] = \{\, |z-1|^{1/2} e^{i\theta/2} \;:\; \theta \in [\![\arg(z-1)]\!] \,\}$ has a branch point at 1.
- i and $-$i are branch points for

$$[\![(z^2+1)^{1/3}]\!] = \Big\{\, |z^2+1|^{1/3} e^{i(\theta+\phi)/3} \;:\; \theta \in [\![\arg(z-i)]\!],\ \phi \in [\![\arg(z+i)]\!] \,\Big\}.$$

- 1 and $-1$ are branch points for

$$[\![\log\big((z-1)/(z+1)\big)]\!] = \{\log|(z-1)/(z+1)| + \mathrm{i}(\theta-\phi) : \theta \in [\![\arg(z-1)]\!],\ \phi \in [\![\arg(z+1)]\!]\}.$$

More generally, for any non-constant rational function $p(z)/q(z)$, the multifunctions $[\![\log\big(p(z)/q(z)\big)]\!]$ and $[\![\big(p(z)/q(z)\big)^\alpha]\!]$ $(\alpha \notin \mathbb{Z})$ have branch points at the roots of $p$ and of $q$ (we assume here that the polynomials have no zeros in common).

Suppose we are given a multifunction $[\![w(z)]\!]$. Our goal is to select a value $f(z) \in [\![w(z)]\!]$, for each $z$ in as large a domain as possible, so that $f$ is holomorphic. In particular, $f$ has to be continuous. We now introduce multibranches. These provide a stepping stone on the way to our goal.

**9.5 Multibranches.** There is a sense in which we can make continuous selections from multifunctions in a natural way. The key idea is the following. Rather than considering $z$ as our variable we introduce, for each branch point $a$, new variables $(r,\theta)$, where $z = a + r\mathrm{e}^{\mathrm{i}\theta}$. We first illustrate how this works in the simplest cases.

For the logarithm

$$[\![\log z]\!] = \{\log r + \mathrm{i}\theta : z = r\mathrm{e}^{\mathrm{i}\theta},\ \theta \in [\![\arg z]\!]\} \quad (0 \neq z),$$

we have multibranches

$$F_k(r,\theta) := \log r + \mathrm{i}(\theta + 2k\pi) \quad (k \in \mathbb{Z}).$$

Each $F_k$ is a continuous function of the polar variables $(r,\theta)$ and $F_k(r,\theta) \in [\![\log z]\!]$ for $0 \neq z = r\mathrm{e}^{\mathrm{i}\theta}$. Furthermore, for any fixed $c \in \mathbb{R}$, and for $0 \neq z = r\mathrm{e}^{\mathrm{i}\theta}$,

$$[\![\log z]\!] = \{F_k(r,\theta) : k \in \mathbb{Z},\ \theta \in [c, c+2\pi)\},$$

with no values repeated, and similarly if $\theta$ is restricted to any interval $(c, c+2\pi]$. We call the set of functions $\{F_k\}_{k\in\mathbb{Z}}$ a **complete set of multibranches** for $[\![\log z]\!]$. We can view these functions as separating the values of the multifunction $[\![\log z]\!]$ into continuous strands.

Now consider what happens when $z$ traces the image of a circle $\gamma(0;r)$. We have $z = r\mathrm{e}^{\mathrm{i}t}$, with $t$ increasing from an initial value, 0, to a final value, $2\pi$. For each $k$, following the continuous multibranch $F_k$ we have

$$[F_k(r,t)]_{t=0} = \log r + 2k\pi\mathrm{i},$$
$$[F_k(r,t)]_{t=2\pi} = \log r + (2k\pi + 2\pi)\mathrm{i} = \log r + 2(k+1)\pi\mathrm{i} = [F_{k+1}(r,t)]_{t=0}.$$

That is, by allowing $z$ to travel anticlockwise round a contour which encloses 0, we transfer from $F_k$ to $F_{k+1}$ in a continuous fashion.

Similarly, consider the square root

$$[\![ z^{1/2} ]\!] = \{ r^{1/2}e^{i\theta/2} \,:\, \theta \in [\![ \arg z ]\!] \}.$$

Then the functions

$$F_+(r,\theta) := r^{1/2}e^{i\theta/2} \quad \text{and} \quad F_-(r,\theta) := r^{1/2}e^{i(\theta+2\pi)/2} = -r^{1/2}e^{i\theta}$$

are continuous functions of $(r,\theta)$. In addition, we have $F_\pm(r,\theta) \in [\![ z^{1/2} ]\!]$ whenever $0 \neq z = re^{i\theta}$. These two functions form a complete set of multibranches for the square root. This time, tracing the image of $\gamma(0;r)$,

$$\begin{aligned} &[F_+(r,t)]_{t=0} = r^{1/2}, \quad [F_+(r,t)]_{t=2\pi} = r^{1/2}e^{2\pi i/2} = [F_-(r,t)]_{t=0}\,; \\ &[F_-(r,t)]_{t=0} = -r^{1/2}, \quad [F_-(r,t)]_{t=2\pi} = -r^{1/2}e^{2\pi i/2} = [F_+(r,t)]_{t=0}\,. \end{aligned}$$

Thus, by letting $z$ encircle 0, we interchange $F_+$ and $F_-$.

In summary, in both these examples, encirclement of a branch point induces a permutation of the multibranches. We may view this permutation as witnessing the nature of the many-valuedness.

For functions with two branch points, say at $a$ and $b$ in $\mathbb{C}$, we need variables $(r,R,\theta,\phi)$, where $z = a + re^{i\theta}$ and $z = b + Re^{i\phi}$. For example, a complete set of multibranches for $[\![ \big((z-a)/(z-b)\big)^{1/3} ]\!]$ would be the functions

$$\begin{aligned} F_1(r,R,\theta,\phi) &= (r/R)^{1/3}e^{i(\theta-phi)/3}, \\ F_2(r,R,\theta,\phi) &= (r/R)^{1/3}e^{2\pi i/3}e^{i(\theta-\phi)/3}, \\ F_3(r,R,\theta,\phi) &= (r/R)^{1/3}e^{4\pi i/3}e^{i(\theta-\phi)/3}. \end{aligned}$$

It is more complicated this time to work out what happens to the multibranches when we encircle one or both of the branch points. So our next task is to explore further the notion of encirclement.

**9.6 Circuits.** We want to track the variation of $\theta$ in $z - a = |z-a|\,e^{i\theta}$ as $z$ moves, with $z \neq a$. To do this tracking in a controlled way we let $z$ trace images of closed paths. Specifically, we let $\gamma$ be a positively oriented contour, where $\gamma$ has parameter interval $[\alpha,\beta]$ (so $\gamma(\alpha) = \gamma(\beta)$). We say $z$ **performs a circuit round** $\gamma$ if we allow $z = \gamma(t)$ to vary with $t$ increasing from the start value $\alpha$ to the final value $\beta$. Basic track readers should take the following as a fact [an advanced track explanation in terms of index is given in 12.11]: as $z = a + e^{i\theta}$ performs a circuit round $\gamma$,

$$\text{if } a \in \mathbf{I}(\gamma), \text{ then } \theta \text{ increases by } 2\pi,$$
$$\text{if } a \in \mathbf{O}(\gamma), \text{ then } \theta \text{ returns to its initial value } 2\pi.$$

Suppose we have a complete set of multibranches, $\{F_\lambda\}_{\lambda\in\Lambda}$, for the multifunction $[\![ w(z) ]\!]$; here $\Lambda$ is some finite or infinite index set. Suppose that $a_1, \ldots, a_n$ are the branch points in $\mathbb{C}$ and let $(r_k, \theta_k)$ be polar variables relative to the point $a_k$. Let $\gamma$ be a positively oriented contour passing through none of $a_1, \ldots, a_n$. For $k = 1, \ldots, n$, let

$$\Theta_k := \begin{cases} \theta_k + 2\pi & \text{if } a_k \in \mathbf{I}(\gamma), \\ \theta_k & \text{if } a_k \in \mathbf{O}(\gamma). \end{cases}$$

We write $F_\lambda \underset{\gamma}{\longrightarrow} F_\mu$ if

$$F_\mu(r_k, \theta_k) = F_\lambda(r_k, \Theta_k) \quad \text{for } k = 1, \ldots, n.$$

We shall say that $\gamma$ is an **admissible contour** for $[\![ w(z) ]\!]$ if $F_\lambda \underset{\gamma}{\longrightarrow} F_\lambda$ for all $\lambda \in \Lambda$; otherwise it is **inadmissible.**

**9.7 Examples (performing circuits).**

- Let $\{F_k\}_{k\in\mathbb{Z}}$ be the multibranches for $[\![ \log z ]\!]$, as defined in 9.5. Then, for all $k \in \mathbb{Z}$,

$$F_k \underset{\gamma}{\longrightarrow} F_\ell, \quad \text{where} \quad F_\ell = \begin{cases} F_{k+1} & \text{if } 0 \in \mathbf{I}(\gamma), \\ F_k & \text{if } 0 \in \mathbf{O}(\gamma). \end{cases}$$

  So $\gamma$ is admissible for $[\![ \log z ]\!]$ if and only if $\gamma$ does not enclose 0.
- For the $n$th root $[\![ z^{1/n} ]\!]$, we have a complete set of multibranches

$$F_k(r, \theta) = e^{2k\pi i/n} r^{1/n} e^{i\theta/n} \quad (k = 0, 1, \ldots, n-1).$$

  We have $F_k(r, \theta + 2\pi) = F_{k+1\,(\mathrm{mod}\,n)}(r, \theta) \quad (k = 0, 1, \ldots, n-1)$, so that $\gamma$ induces a cyclic permutation

$$F_0 \underset{\gamma}{\longrightarrow} F_1 \underset{\gamma}{\longrightarrow} \ldots \underset{\gamma}{\longrightarrow} F_{n-1} \underset{\gamma}{\longrightarrow} F_0$$

  of the multibranches if $0 \in \mathbf{I}(\gamma)$. If 0 is outside $\gamma$ then all multibranches are unchanged after a circuit round $\gamma$.

- Consider the multifunction $[\![ w(z) ]\!] = [\![ ((z-a)(z-b))^{1/2} ]\!]$ $(a, b \in \mathbb{C}, a \neq b)$. We have branch points at $a$ and $b$. Write $z - a = re^{i\theta}$ and $z - b = Re^{i\phi}$. We take multibranches

$$F_+(r, R, \theta, \phi) = (rR)^{1/2}e^{i(\theta+\phi)/2} \text{ and } F_-(r, R, \theta, \phi) = -(rR)^{1/2}e^{i(\theta+\phi)/2}.$$

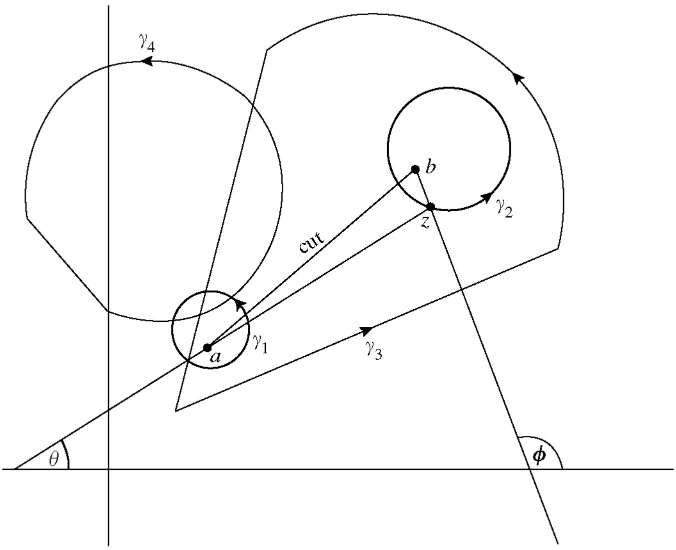



**Figure 9.1** Circuits for $[\![ ((z-a)(z-b))^{1/2} ]\!]$

We show in Table 9.1 the effect on the angles $\theta$ and $\phi$ and on the multibranches of performing circuits round contours $\gamma_1, \gamma_2, \gamma_3$, and $\gamma_4$, as shown in Fig. 9.1. We consider this example further in 9.11.

## Cuts and holomorphic branches

In the previous section we identified argument multifunctions as the root cause of many-valuedness of logarithms and non-integer powers of rational functions. We introduced multibranches and circuits as tools for analysing the effect of this many-valuedness. We now show how to remedy many-valuedness by working in a plane cut so that inadmissible contours are outlawed. We then make selections of values of multifunctions which are linked to the arg restrictions imposed by the cut(s).

**Table 9.1** Circuits for $[\![\, ((z-a)(z-b))^{1/2} \,]\!]$

| | $\gamma_1$ | $\gamma_2$ | $\gamma_3$ | $\gamma_4$ |
|---|---|---|---|---|
| $\theta \uparrow$ | $2\pi$ | $0$ | $2\pi$ | $0$ |
| $\phi \uparrow$ | $0$ | $2\pi$ | $0$ | $2\pi$ |
| $\frac{1}{2}(\theta+\phi) \uparrow$ | $\pi$ | $\pi$ | $2\pi$ | $0$ |
| | $F_+ \underset{\gamma}{\longleftrightarrow} F_-$ | $F_+ \underset{\gamma}{\longleftrightarrow} F_-$ | no change | no change |

**9.8 Cutting the plane.** Suppose that we have a multifunction $[\![\, w(z) \,]\!]$ for which we have identified the branch points and a complete set of multibranches. Suppose also that, by considering circuits round contours which include or exclude the various branch points, we have found which contours are admissible and which are inadmissible. We now wish to restrict movement of $z$ so that inadmissible contours are outlawed. We do this by means of cuts in the plane which we forbid $z$ to cross. For example, a cut along $[0,\infty)$ would outlaw any $\gamma$ enclosing 0, but would not outlaw $\gamma(-2;1)$. A cut along $[-\mathrm{i},\mathrm{i}]$ would outlaw any contour enclosing one but not both of i and $-$i.

For our purposes it will be sufficient to consider cuts of the following forms:

- along an infinite ray from a branch point $a$;
- along a line segment $[a,b]$ joining branch points $a$ and $b$.

We do not remove points of a cut from the plane, but we do think of a cut as having two edges.

Consider a cut of either type with $a$ as an endpoint and let $z-a=r\mathrm{e}^{\mathrm{i}\theta}$. On the cut, $\theta=\alpha$, where $\alpha$ is a fixed constant, determined up to an integer multiple of $2\pi$. For any point $z$ not on the cut, we can specify a unique value of $\theta$ by requiring that $\theta\in(\alpha,\alpha+2\pi)$. For points on the two edges of the cut (excluding $a$), $\theta=\alpha$ defines one edge and $\theta=\alpha+2\pi$ the other.

**9.9 Examples (plane-cutting).** In the following examples, notice how the restrictions on the angles tie up with the positions of the cuts.

- In the plane cut along $\{\, \mathrm{i}y : y \geqslant 1 \,\}$, any point other than i may be uniquely specified as $z=|z|\mathrm{e}^{\mathrm{i}\theta}$, where $\pi/2<\theta\leqslant 5\pi/2$.

- Cut the plane along $[-1,1]$. For any point $z$ other than 1 or $-1$,

$$z-1=|z-1|\,\mathrm{e}^{\mathrm{i}\theta} \quad (-\pi<\theta\leqslant\pi),$$
$$z+1=|z+1|\,\mathrm{e}^{\mathrm{i}\varphi} \quad (0\leqslant\varphi<2\pi),$$

with $\theta$ and $\varphi$ uniquely determined. We have used 'upper edge' angle values for points on the cut. See Fig. 9.2. We could alternatively have taken 'lower edge' values, but not a mixture. So, for example, upper-edge values for $z-1$ and lower edge values for $z+1$ would not be allowed.

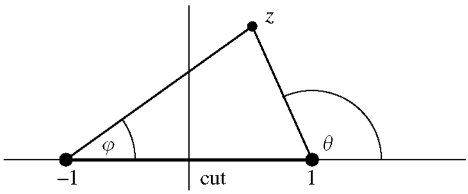



**Figure 9.2** Plane cut along $[-1,1]$

All the machinery is now in place for us to extract holomorphic functions from multifunctions.

**9.10 From multibranches to holomorphic branches.** Let $[\![\, w(z) \,]\!]$ be a multifunction with branch points $a_1,\ldots,a_n$ in $\mathbb{C}$. Suppose that $\{F_\lambda\}_{\lambda\in\Lambda}$ is a complete set of multibranches and suppose that we have cut the plane so as to forbid inadmissible contours. In one sense we eliminate many-valuedness by switching from the variable $z$ to the polar variables $(r_1,\ldots,r_n,\theta_1,\ldots,\theta_n)$ used with multibranches, where $z=a_m+r_m\mathrm{e}^{\mathrm{i}\theta_m}$ $(m=1,\ldots,n)$. Our aim is to impose arg restrictions as dictated by the cuts, to enable us to revert to the variable $z$ so that

$$f_\lambda(z):=F_\lambda(r_1,\ldots,r_n,\theta_1,\ldots,\theta_n)$$

is uniquely determined by $z$. If $f_\lambda$ is holomorphic in the plane with the points of the cut(s) removed, then it is called a **holomorphic branch.** The holomorphic branches we defined in 7.16 and 7.17 are instances.

Here are two examples of taming multifunctions which have more than one branch point in $\mathbb{C}$.

**9.11 Examples (holomorphic branches).**

- Consider $[\![ \big((z-1)(z+1)\big)^{1/2} ]\!]$. We showed in 9.6 that we must outlaw contours which encircle one, but not both, of the branch points 1 and $-1$. We achieve this by cutting the plane along $[-1,1]$ (see 9.8). In the cut plane we then have two holomorphic branches

$$\pm \left|z^2-1\right|^{1/2} \mathrm{e}^{\mathrm{i}(\theta+\varphi)/2},$$
$$\theta \in [\![ \arg(z-1) ]\!],\ -\pi < \theta \leqslant \pi,\quad \varphi \in [\![ \arg(z+1) ]\!],\ 0 \leqslant \varphi < 2\pi.$$

- Consider $[\![ \log(z^2-1) ]\!]$. There are branch points at $\pm 1$. We define multibranches

$$G_k(r,R,\theta,\varphi) = \log\left|z^2-1\right| + \mathrm{i}(\theta+\varphi+2k\pi) \quad (k \in \mathbb{Z}).$$

A circuit round a contour $\gamma$ increases $\theta+\varphi$ by $2\pi$ if $\gamma$ encloses just one of the points 1 and $-1$ and increases $\theta+\varphi$ by $4\pi$ if $\gamma$ encloses both 1 and $-1$. If $\gamma$ encloses neither 1 nor $-1$, then $\theta+\varphi$ is not changed. Therefore we must outlaw all contours which enclose either, or both, of the branch points. This can be achieved by cuts along $(-\infty,-1]$ and $[1,\infty)$ (see Fig. 9.3). Contrast this with the preceding example.

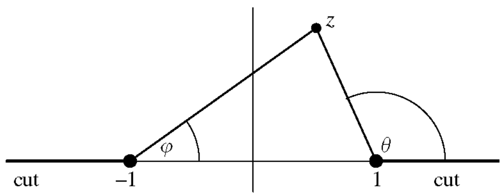



**Figure 9.3** Cut plane for $[\![ \log(z^2-1) ]\!]$

Holomorphic branches are given, for $k \in \mathbb{Z}$, by

$$g_k(z) = \log\left|z^2-1\right| + \mathrm{i}(\theta+\varphi+2k\pi), \quad \text{where}$$
$$z-1 = |z-1|\,\mathrm{e}^{\mathrm{i}\theta},\ 0 \leqslant \theta < 2\pi,\quad z+1 = |z+1|\,\mathrm{e}^{\mathrm{i}\varphi},\ -\pi < \varphi \leqslant \pi.$$

**9.12 A secret revealed: branch points at $\infty$.** The difference between the two examples in 9.11 may seem mysterious, as may the non-uniqueness of cuts which are infinite rays. The mystery disappears once we investigate the behaviour of multifunctions at $\infty \in \widetilde{\mathbb{C}}$. We say that $\infty$ is a branch point of $[\![\, w(z) \,]\!]$ if 0 is a branch point of $\widetilde{w}$, where $[\![\, \widetilde{w}(\zeta) \,]\!] := [\![\, w(1/\zeta) \,]\!]$ (for $\zeta \neq 0$ and $1/\zeta$ in the domain of definition of $w$).

Circuits round 0 in the $\zeta$-plane, where $\zeta = 1/z$, must be outlawed if 0 is a branch point of $\widetilde{w}$. Back in the $z$-plane, we must cut along at least one infinite ray if $\infty$ is a branch point. Remember that any infinite ray $\arg(z - a) = \mu$ joins $a$ to $\infty$. Table 9.2 gives some examples.

**Table 9.2** Examples of cuts

| $[\![\, w(z) \,]\!]$ | $[\![\, \widetilde{w}(z) \,]\!]$ | branch points of $w$ | possible cuts |
|---|---|---|---|
| $[\![\, z^{1/2} \,]\!]$ | $[\![\, z^{-1/2} \,]\!]$ | $0, \infty$ | $[0, \infty)$ or $(-\infty, 0]$ |
| $[\![\, \log(z^{-1}) \,]\!]$ | $[\![\, \log z \,]\!]$ | $0, \infty$ | $[0, \infty)$ or $(-\infty, 0]$ |
| $[\![\, (z^2 - 1)^{\frac{1}{2}} \,]\!]$ | $[\![\, (1 - z^2)^{1/2}/z \,]\!]$ | $\pm 1$ | $[-1, 1]$ |
| $\left[\!\!\left[ \left(\dfrac{z-1}{z+1}\right)^{\frac{1}{2}} \right]\!\!\right]$ | $\left[\!\!\left[ \left(\dfrac{1-z}{1+z}\right)^{1/2} \right]\!\!\right]$ | $\pm 1$ | $[-1, 1]$ |
| $\left[\!\!\left[ \log\left(\dfrac{z-1}{z+1}\right) \right]\!\!\right]$ | $\left[\!\!\left[ \log\left(\dfrac{1-z}{1+z}\right) \right]\!\!\right]$ | $\pm 1$ | $[-1, 1]$ |
| $[\![\, \log(z^2 - 1) \,]\!]$ | $[\![\, \log((1 - z^2)/z^2) \,]\!]$ | $\pm 1, \infty$ | $(-\infty, -1]$ & $[1, \infty)$ |

**9.13 Summing up.** Given a multifunction $w$, we seek to

(1) define a complete set of multibranches,

(2) locate the branch points (in $\widetilde{\mathbb{C}}$),

(3) identify which contours are admissible and which are not,

(4) cut the plane between pairs of branch points so as to debar inadmissible contours, making no more cuts than are necessary,

and so to

(5) specify holomorphic branches.

Here (4) serves as a device for stipulating the arg restrictions in a way which makes them transparent. In simple cases it may be possible to bypass some of the steps and to define holomorphic branches directly.

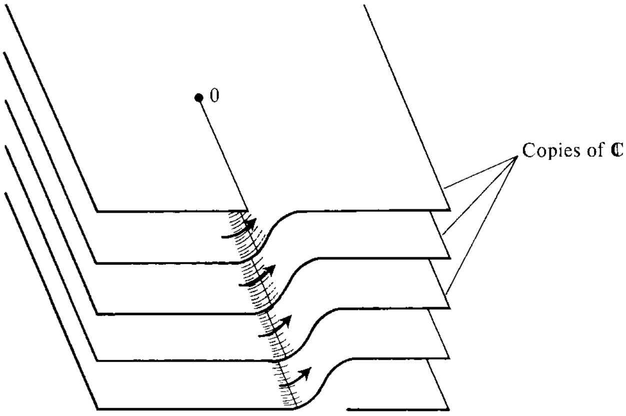



**Figure 9.4** Riemann surface for $[\![\log z]\!]$

**9.14 Riemann surfaces.** It could be argued that we have attacked the many-valuedness problem in an ostrich-like way. Maybe, instead of working with an individual branch of a multifunction in a cut plane, it would be better to keep all branches in play, with one copy of the plane on which to define each. In other words, we treat the aggregate of branches as a single function on a domain set consisting of many copies of the plane. These copies are glued together so that in moving from one to another we pass continuously from one branch of the multifunction to another. The resulting structure of 'parallel universes' is known as a Riemann surface for the multifunction. A multi-storey car park provides a good mental picture. The floors of the car park represent copies of the plane and the ramps taking cars up and down between levels indicate how these copies are pasted together. The Riemann surface for the logarithm is modelled by a car park with infinitely many floors, each of infinite extent, with a semi-infinite ramp joining each floor to the next; see Fig. 9.4. For more complicated multifunctions, the car park designer might be said to have a warped sense of humour.

An extensive theory of Riemann surfaces exists and provides a framework for an advanced study of multifunctions affording, in return for some sophisticated analysis, far greater insights than the naive plane-cutting approach. However, for the sort of multifunction problems we discuss in this book, cut planes serve quite adequately.

## Exercises

9.1 Let $\gamma$ be the circular path $\gamma(0;2)$. Suppose values of $f(z) \in [\![ w(z) ]\!]$ are selected so that $f(\gamma(t))$ varies continuously as $t$ increases from 0 to $2\pi$, with $f(\gamma(0))$ real. Determine the initial value $f(\gamma(0))$ and final value $f(\gamma(2\pi))$ when $[\![ w(z) ]\!]$ is

$$\text{(i) } [\![ \log z ]\!], \quad \text{(ii) } [\![ \log(z^{-1}) ]\!], \quad \text{(iii) } [\![ \log(1+z) ]\!], \quad \text{(iv) } [\![ \log(z^2) ]\!].$$

9.2 Repeat Exercise 9.1 for the following multifunctions:

$$\text{(i) } [\![ (z-1)^{1/3} ]\!], \quad \text{(ii) } [\![ (z^2+1)^{3/2} ]\!], \quad \text{(iii) } [\![ z^{\pi/3} ]\!].$$

(Assume that $f(\gamma(0))$ is real and positive.)

9.3 Verify that the multifunctions below have branch points as indicated and that the cuts suggested outlaw precisely the inadmissible contours. In each case, specify a holomorphic branch.

(i) $[\![ (z-\mathrm{i})^{1/2} ]\!]$ (branch points $\mathrm{i}, \infty$; cut along $\{ \mathrm{i}y : y \geqslant 1 \}$).

(ii) $[\![ \big((z-1)/(z+1)\big)^{3/4} ]\!]$ (branch points $\pm 1$; cut along $[-1,1]$).

(iii) $[\![ \big(z(z-1)\big)^{-1/2} ]\!]$ (branch points $0, 1$; cut along $[0,1]$).

(iv) $[\![ \big(z(z-1)\big)^{2/3} ]\!]$ (branch points $0, 1, \infty$; cuts along $(-\infty,-1]\&[1,\infty)$).

(v) $[\![ \log(z^2+1) ]\!]$ (branch points $\pm\mathrm{i}$; cuts along $\{ \mathrm{i}y : |y| \geqslant 1 \}$).

9.4 For each of the following multifunctions, locate the branch points (in $\widetilde{\mathbb{C}}$), suggest how the plane should be cut, and specify a holomorphic branch.

$$\text{(i) } [\![ (z^2-1)^{-1/2} ]\!], \quad \text{(ii) } [\![ \big(z^2(1-z)\big)^{1/2} ]\!], \quad \text{(iii) } [\![ (z+z^{-1})^{1/2} ]\!].$$

9.5 For each of the following multifunctions, locate the branch points (in $\widetilde{\mathbb{C}}$), specify multibranches, suggest how the plane should be cut, and specify a holomorphic branch.

$$\text{(i) } [\![ z^{1/2} \log z ]\!], \quad \text{(ii) } [\![ \log z^{1/2} ]\!], \quad \text{(iii) } [\![ \big((z-1)(z-\omega)(z-\omega^2)\big)^{1/2} ]\!],$$

where $\omega = \mathrm{e}^{2\pi\mathrm{i}/3}$.

9.6 Let $a_1, \ldots, a_n$ be distinct points of $\mathbb{C}$. Prove that the multifunction

$$[\![ \big((z-a_1)\ldots(z-a_n)\big)^{1/2} ]\!]$$

has branch points at each point $a_m$ and has a branch point at $\infty$ if $n$ is odd but not if $n$ is even.

# 10 Integration in the complex plane

This chapter continues the study of paths begun in Chapter 4. It contains basic material on integrals along paths which underpins the presentation of Cauchy's theorem in Chapters 11 and 12.

## Integration along paths

**10.1 Paths: a recap.** For the definitions of a path and of concepts relating to paths, consult 4.2. Recall in particular that a path is defined by a function $\gamma\colon [\alpha, \beta] \to \mathbb{C}$ and that $\gamma^*$ denotes the image traced out by $\gamma(t)$ (for $\alpha \leqslant t \leqslant \beta$). To qualify as a path, the function $\gamma$ is required to satisfy a rather technical differentiability condition. Intuitively, we may interpret this condition as telling us that the image $\gamma^*$ is made up of finitely many smooth sections (as the image of a contour certainly is).

Essentially, the definition of a path is set up so that we are able to apply the theory of integration of continuous functions, piecewise, to the finitely many smooth paths which join to form $\gamma$.

**10.2 Integration of real- and complex-valued functions.** We assume familiarity with integration of real-valued functions on compact intervals in $\mathbb{R}$, at least at a fairly basic level. Specifically, we take for granted simple techniques for evaluating real integrals, elementary properties of integrals (such as linearity) and the fact that, at the least, continuous real-valued functions on compact intervals in $\mathbb{R}$ are integrable. [Here 'integrable' may be interpreted as having the meaning it has in any treatment of basic integration theory, whether a Riemann-style approach or a more sophisticated one leading to the Lebesgue integral.]

Continuous functions are not quite adequate for our needs. We say that a (real- or complex-valued) function $h$ is **piecewise continuous** on a compact interval $[\alpha, \beta]$ in $\mathbb{R}$ if there exist points $\alpha = t_0 < t_1 < \cdots < t_n = \beta$ and continuous functions $h_k$ on $[t_k, t_{k+1}]$ such that $h(t) = h_k(t)$ for $t \in (t_k, t_{k+1})$ $(k = 0, \ldots, n-1)$; $h$ need not be defined at any or at some of the poimts $t_k$.

Essentially, the definition means that $h$ is continuous except, possibly, for a finite number of jump discontinuities. A real-valued piecewise continuous function is integrable, with

$$\int_\alpha^\beta h(t)\,\mathrm{d}t = \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}} h_k(t)\,\mathrm{d}t.$$

Let $g$ be a complex-valued function (defined on $[\alpha,\beta] \subseteq \mathbb{R}$) and write $g$ as $\operatorname{Re} g + \mathrm{i}\operatorname{Im} g$, where $\operatorname{Re} g, \operatorname{Im} g\colon [\alpha,\beta] \to \mathbb{R}$. We say $g$ is **integrable** if and only if $\operatorname{Re} g$ and $\operatorname{Im} g$ are both integrable and in that case we define

$$\int_\alpha^\beta g(t)\,\mathrm{d}t := \int_\alpha^\beta \operatorname{Re} g(t)\,\mathrm{d}t + \mathrm{i}\int_\alpha^\beta \operatorname{Im} g(t)\,\mathrm{d}t.$$

For example,

$$\int_0^{2\pi} \mathrm{e}^{\mathrm{i}t}\,\mathrm{d}t := \int_0^{2\pi} \cos t\,\mathrm{d}t + \mathrm{i}\int_0^{2\pi} \sin t\,\mathrm{d}t = \big[\sin t\big]_0^{2\pi} + \mathrm{i}\big[-\cos t\big]_0^{2\pi} = 0.$$

In the sequel, where we manipulate complex-valued integrals without comment, we are using properties which carry over easily from the real-valued case; examples are linearity and the theorem on substitution used in 10.5. These properties and their derivations can be found, for example, in **[6]** and **[2]**.

**10.3 The integral of a function along a path.** Let $\gamma$ be a path with parameter interval $[\alpha,\beta]$. There exist points $\alpha = t_0 < t_1 < \cdots < t_n = \beta$ such that $\gamma$ restricted to each $[t_k, t_{k+1}]$ coincides with a continuously differentiable function on $[t_k, t_{k+1}]$. At the intermediate points $t_k$, $\gamma'$ may not exist. Let $f\colon \gamma^* \to \mathbb{C}$ be continuous. We define

$$\int_\gamma f(z)\,\mathrm{d}z := \int_\alpha^\beta f(\gamma(t))\gamma'(t)\,\mathrm{d}t$$

and call this the **integral of** $f$ **along** $\gamma$, or **round** $\gamma$ if $\gamma$ is closed. The integral on the right-hand side makes sense because $(f\circ\gamma)\gamma'\colon t \mapsto f(\gamma(t))\gamma'(t)$ is piecewise continuous, and hence integrable. We can motivate the definition on a formal, symbol-juggling, level by replacing $z$ by $\gamma(t)$ and $\mathrm{d}z$ by $\gamma'(t)\mathrm{d}t$.

[In vector calculus, integrals along paths in $\mathbb{R}^2$ are important. These are known as **line integrals** or **curvilinear integrals** and are usually defined, Riemann-fashion, by breaking up the parameter interval of a path given by $\gamma(t) = (x(t), y(t))$ into small subintervals and forming approximating sums.

Given $f = u + \mathrm{i}v$, the formalism

$$\begin{aligned}\int_\gamma f(z)\,\mathrm{d}z &= \int_\alpha^\beta (u+\mathrm{i}v)(x'+\mathrm{i}y')\,\mathrm{d}t \\ &= \int_\alpha^\beta (ux'-vy')\,\mathrm{d}t + \mathrm{i}\int_\alpha^\beta (uy'+vx')\,\mathrm{d}t \\ &= \int_\gamma u\,\mathrm{d}x - \int_\gamma v\,\mathrm{d}y + \mathrm{i}\int_\gamma v\,\mathrm{d}x + \mathrm{i}\int_\gamma u\,\mathrm{d}y\end{aligned}$$

correctly relates line integrals of $u(x(t), y(t))$ and $v(x(t), y(t))$ along $\gamma$ to the complex integral $\int_\gamma f(z)\,\mathrm{d}z$.]

The following simple consequence of the fundamental identity $\mathrm{e}^{2k\pi\mathrm{i}} = 1$ ($k \in \mathbb{Z}$) will be used again and again in later chapters.

**10.4 Example: the Fundamental integral.** For $a \in \mathbb{C}$ and $r > 0$,

$$\int_{\gamma(a;r)} (z-a)^n\,\mathrm{d}z = \begin{cases} 0 & (n \neq -1), \\ 2\pi\mathrm{i} & (n = -1), \end{cases}$$

where $\gamma(a;r)$ denotes the circular contour having centre $a$ and radius $r$ (recall 4.3).

**Proof** Since $\gamma(a;r)(t) = a + r\mathrm{e}^{\mathrm{i}t}$ ($t \in [0, 2\pi]$), the definitions in 10.3 and 10.2 give

$$\begin{aligned}\int_{\gamma(a;r)} (z-a)^n\,\mathrm{d}z &= \int_0^{2\pi} (r\mathrm{e}^{\mathrm{i}t})^n r\mathrm{i}\mathrm{e}^{\mathrm{i}t}\,\mathrm{d}t \\ &= \mathrm{i}r^{n+1}\int_0^{2\pi} \mathrm{e}^{\mathrm{i}(n+1)t}\,\mathrm{d}t \\ &= \mathrm{i}r^{n+1}\left(\int_0^{2\pi} \cos(n+1)t\,\mathrm{d}t + \mathrm{i}\int_0^{2\pi} \sin(n+1)t\,\mathrm{d}t\right) \\ &= \begin{cases} \mathrm{i}r^{n+1}\left(\left[\dfrac{\sin(n+1)t}{n+1}\right]_0^{2\pi} - \mathrm{i}\left[\dfrac{\cos(n+1)t}{n+1}\right]_0^{2\pi}\right) & (n \neq -1), \\ \mathrm{i}[1]_0^{2\pi} & (n = -1), \end{cases}\end{aligned}$$

from which the conclusion follows immediately. □

We now clear up some technical points. The results, which seem quite transparent in particular cases, are more important than the details of the proofs.

**10.5 Technical lemma (integrals along paths).** Suppose that $\gamma$ is a path with parameter interval $[\alpha, \beta]$ and that $f\colon \gamma^* \to \mathbb{C}$ is continuous.

(1) **Reversal** $\int_{-\gamma} f(z)\,\mathrm{d}z = -\int_\gamma f(z)\,\mathrm{d}z$.

(2) **Joining** Let $\alpha < \tau < \beta$ and let $\gamma_1 = \gamma\restriction[\alpha,\tau]$ and $\gamma_2 = \gamma\restriction[\tau,\beta]$. Then

$$\int_\gamma f(z)\,\mathrm{d}z = \int_{\gamma_1} f(z)\,\mathrm{d}z + \int_{\gamma_2} f(z)\,\mathrm{d}z.$$

(3) **Reparametrization** Let $\widetilde{\gamma}$ be another path, with parameter interval $[\widetilde{\alpha}, \widetilde{\beta}]$ and suppose that $\widetilde{\gamma} = \gamma \circ \psi$, where $\psi$ is a function which maps $[\widetilde{\alpha}, \widetilde{\beta}]$ onto $[\alpha, \beta]$ and has a positive continuous derivative. Then

$$\int_{\widetilde{\gamma}} f(z)\,\mathrm{d}z = \int_\gamma f(z)\,\mathrm{d}z.$$

**Proof** (1) and (2) are easily deduced from the definitions. In proving (3) we may, thanks to (2), assume that $\gamma$, and hence also $\widetilde{\gamma}$, is smooth. For $t \in [\widetilde{\alpha}, \widetilde{\beta}]$,

$$\widetilde{\gamma}'(t) = \gamma'(\psi(t))\psi'(t)$$

(this is the Chain rule in a real/complex hybrid form). Making the substitution $s = \psi(t)$ is legitimate because of the hypotheses on $\psi$ and so

$$\begin{aligned}\int_{\widetilde{\gamma}} f(z)\,\mathrm{d}z &= \int_{\widetilde{\alpha}}^{\widetilde{\beta}} f(\widetilde{\gamma}(t))\widetilde{\gamma}'(t)\,\mathrm{d}t \\ &= \int_{\widetilde{\alpha}}^{\widetilde{\beta}} f(\gamma(\psi(t)))\gamma'(\psi(t))\psi'(t)\,\mathrm{d}t \\ &= \int_\alpha^\beta f(\gamma(s))\gamma'(s)\,\mathrm{d}s \\ &= \int_\gamma f(z)\,\mathrm{d}z. \qquad \square\end{aligned}$$

The reparametrization result 10.5(3) tells us that under quite mild conditions $\int_\gamma f(z)\,\mathrm{d}z$ depends only on the image $\gamma^*$ and on the direction in which it is traced, and not on the parametrization chosen. In particular, we can translate and rescale the parameter interval. Translation is needed in the proof of the following corollary to 10.5(2); recall the definition of join in 4.2.

**10.6 Proposition (integral along a join of paths).** Suppose that $\gamma$ is a path with parameter interval $[\alpha, \beta]$ and is the join of paths $\gamma_1, \gamma_2, \ldots, \gamma_n$ and let $f\colon \gamma^* \to \mathbb{C}$ be continuous. Then

$$\int_\gamma f(z)\,\mathrm{d}z = \sum_{k=1}^{n} \int_{\gamma_k} f(z)\,\mathrm{d}z.$$

**10.7 Example (10.5 and 10.6 in action).** We compute $\int_\gamma z^2\,\mathrm{d}z$, where $\gamma$ is the semicircular contour formed by joining $\gamma_1 := [-R, R]$ and $\gamma_2 := \Gamma_R$; see Fig. 10.1. We have

$$\gamma_1(t) = (1-t)(-R) + tR \quad \text{and} \quad \gamma_1'(t) = 2R \quad (t \in [0,1]),$$
$$\gamma_2(t) = R\mathrm{e}^{\mathrm{i}t} \quad \text{and} \quad \gamma_2'(t) = \mathrm{i}R\mathrm{e}^{\mathrm{i}t} \quad (t \in [0, 2\pi]).$$

Hence

$$\begin{aligned}\int_\gamma z^2\,\mathrm{d}z &= \int_0^1 ((2t-1)R)^2 2R\,\mathrm{d}t + \int_0^\pi R^2\mathrm{e}^{2\mathrm{i}t}\,\mathrm{i}R\mathrm{e}^{\mathrm{i}t}\,\mathrm{d}t \\ &= \Big[2R^3\big(\frac{4}{3}t^3 - 2t^2 + t\big)\Big]_0^1 + \Big[\frac{1}{3}R^3\mathrm{e}^{3\mathrm{i}t}\Big]_0^\pi \\ &= 0.\end{aligned}$$

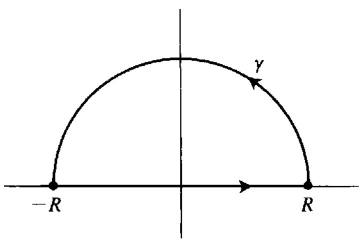



**Figure 10.1** Contour for Example 10.7

## The Fundamental theorem of calculus

The usual way to evaluate real integrals between fixed limits (definite integrals) is to recognize the integrand as a continuous derivative and then to apply the Fundamental theorem of calculus. There is an analogous result for complex integrals. The calculations in 10.7 are special cases of those in the proof below.

**10.8 Fundamental theorem of calculus.** Suppose that $\gamma$ is a path with parameter interval $[\alpha, \beta]$, that $F$ is defined on an open set containing $\gamma^*$, and that $F'(z)$ exists and is continuous at each point of $\gamma^*$. Then

$$\int_\gamma F'(z)\,\mathrm{d}z = \begin{cases} F(\gamma(\beta)) - F(\gamma(\alpha)) & \text{in general,} \\ 0 & \text{if } \gamma \text{ is closed.} \end{cases}$$

**Proof** We first assume that $\gamma$ is smooth. The hypotheses on $F$ are more than strong enough to imply that $F \circ \gamma$ is differentiable on $[\alpha, \beta]$ with $(F \circ \gamma)'(t) = F'(\gamma(t))\gamma'(t)$ by the Chain rule. Then

$$\begin{aligned} \int_\gamma F(z)\,\mathrm{d}z &= \int_\alpha^\beta F'(\gamma(t))\gamma'(t)\,\mathrm{d}t \\ &= \int_\alpha^\beta (F \circ \gamma)'(t)\,\mathrm{d}t \\ &= \int_\alpha^\beta (\mathrm{Re}(F \circ \gamma)'(t) + \mathrm{i}\,\mathrm{Im}(F \circ \gamma)'(t))\,\mathrm{d}t \\ &= \Big[\mathrm{Re}(F \circ \gamma)(t)\Big]_\alpha^\beta + \mathrm{i}\Big[\mathrm{Im}(F \circ \gamma)(t)\Big]_\alpha^\beta \\ &= F(\gamma(\beta)) - F(\gamma(\alpha)). \end{aligned}$$

The penultimate line is obtained by applying the real Fundamental theorem of calculus (see for example [**6**], [**3**], or [**2**]) to $\mathrm{Re}(F \circ \gamma)$ and $\mathrm{Im}(F \circ \gamma)$.

In the general case, we choose $\alpha = t_0 < t_1 < \cdots < t_n = \beta$ such that $\gamma\lceil[t_k, t_{k+1}]$ is smooth, for $k = 0, \ldots, n-1$. By the above and 10.6,

$$\begin{aligned} \int_\gamma F(z)\,\mathrm{d}z &= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} F'(\gamma(t))\gamma'(t)\,\mathrm{d}t \\ &= \sum_{k=0}^{n-1} \big(F(\gamma(t_{k+1})) - F(\gamma(t_k))\big) \\ &= F(\gamma(\beta)) - F(\gamma(\alpha)). \end{aligned}$$

□

**10.9 The status of the complex Fundamental theorem of calculus.** The Fundamental theorem of calculus in complex analysis should be treated, in a way that its real counterpart is not, as an interim result. It is a stepping stone to Cauchy's theorem and its consequences, and these results largely supersede it for both theoretical and computational purposes.

When complex integrals cannot be evaluated explicitly (and sometimes when they can) the following estimate of magnitude is invaluable.

**10.10 Estimation theorem.** Suppose that $\gamma$ is a path with parameter interval $[\alpha, \beta]$ and that $f\colon \gamma^* \to \mathbb{C}$ is continuous. Then

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant \int_\alpha^\beta |f(\gamma(t))\gamma'(t)|\,\mathrm{d}t.$$

In particular, if $|f(z)| \leqslant M$ for all $z \in \gamma^*$ then

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant M \times \text{length}(\gamma),$$

where, by definition,

$$\text{length}(\gamma) := \int_\alpha^\beta |\gamma'(t)|\,\mathrm{d}t.$$

(For line segments and circular arcs, and hence also for contours, this definition gives the value we expect the length to have.)

**10.11 Examples (estimation).** Using estimation correctly depends on familiarity with the inequalities in 1.9.

- Let $f(z) = (z^4 + 1)^{-1}$ and let $\gamma = \Gamma_R$. Then, by definition,

$$\int_\gamma f(z)\,\mathrm{d}z = \int_0^\pi \frac{1}{R^4 \mathrm{e}^{4\mathrm{i}t} + 1} \mathrm{i}R\mathrm{e}^{\mathrm{i}t}\,\mathrm{d}t,$$

  the value of which is not obvious. However we do have, by 10.10 and 1.9(3),

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant \int_0^\pi \left|\frac{R\mathrm{i}\mathrm{e}^{\mathrm{i}t}}{R^4 \mathrm{e}^{4\mathrm{i}t} + 1}\right| \mathrm{d}t \leqslant \frac{R\pi}{|R^4 - 1|}.$$

- Take $f(z) = 1/z$, $\gamma(t) = \mathrm{e}^{\mathrm{i}t}$ $(t = [0, 2\pi])$. Then $|f(\gamma(t))| = 1$ and $|\gamma'(t)| = 1$. The Estimation theorem gives

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant \int_0^{2\pi} 1\,\mathrm{d}t = 2\pi,$$

which is consistent with 10.4. Compare this with the fallacious estimate

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant \int_\gamma |f(z)|\,\mathrm{d}z = \int_\gamma 1\,\mathrm{d}z = \int_0^\pi \mathrm{i}e^{\mathrm{i}t}\,\mathrm{d}t = 0.$$

The error lies in misplaced modulus signs. The moduli must enclose the *entire* parametrized integrand $f(\gamma(t))\gamma'(t)$ and not just $f(\gamma(t)) = f(z)$. A legitimate shorthand for $\int_\gamma |f(\gamma(t))\gamma'(t)|\,\mathrm{d}t$ is $\int_\gamma |f(z)|\,|\mathrm{d}z|$, which must not be confused with $\int_\gamma |f(z)|\,\mathrm{d}z$.

## Exercises

**Exercises from the text.** Verify that the definition given in 10.10 for the length of a path $\gamma$ gives the expected value when $\gamma$ is a line segment or a circular arc.

10.1 Evaluate $\int_\gamma f(z)\,\mathrm{d}z$ when

(i) $f(z) = z^2$, $\gamma(t) = e^{\mathrm{i}t}$ $(t \in [-\pi/2, \pi/2])$,

(ii) $f(z) = \operatorname{Re} z$, $\gamma(t) = t + \mathrm{i}t^2$ $(t \in [0, 1])$,

(iii) $f(z) = 1/z$, $\gamma(t) = e^{-\mathrm{i}t}$ $(t \in [0, 8\pi])$,

(iv) $f(z) = e^z$, $\gamma$ the join of $[0, 1]$, $[1, 1+\mathrm{i}]$, and $[1+\mathrm{i}, \mathrm{i}]$,

(v) $f(z) = |z|^4$, $\gamma = [-1+\mathrm{i}, 1+\mathrm{i}]$.

10.2 Evaluate $\int_{\gamma(0;1)} f(z)\,\mathrm{d}z$ when $f(z)$ is

(i) $|z|^4$, (ii) $(\operatorname{Re} z)^2$, (iii) $z^{-2}(z^4 - 1)$, (iv) $\sin z$.

(Use the Fundamental theorem of calculus where applicable.)

10.3 By integrating $(R+z)/(z(R-z))$ round a suitable contour, prove that

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{R^2 - r^2}{R^2 - 2Rr\cos\theta + r^2}\,\mathrm{d}\theta = 1 \quad (0 \leqslant r < R).$$

(Hint: use partial fractions.)

10.4 Prove the following properties of length($\gamma$), the length of a path $\gamma$.

(i) $\text{length}(-\gamma) = \text{length}(\gamma)$.

(ii) If $\gamma$ is the join of paths $\gamma_1$ and $\gamma_2$ then

$$\text{length}(\gamma) = \text{length}(\gamma_1) + \text{length}(\gamma_2).$$

(iii) If $\tilde{\gamma}$ is obtained from $\gamma$ by reparametrizing $\gamma$ as in 10.5(3), then length$(\gamma)$ = length$(\tilde{\gamma})$.

10.5 Use the Estimation theorem to obtain the following upper bounds:

$$\text{(i)} \left| \int_{\gamma(1;2)} \frac{1}{z}\, \mathrm{d}z \right| \leqslant 4\pi, \qquad \text{(ii)} \left| \int_{\gamma(0;R)} \frac{z-1}{z+1}\, \mathrm{d}z \right| \leqslant \frac{2\pi R(R+1)}{R-1},$$

$$\text{(iii)} \left| \int_{\Gamma_R} \frac{\mathrm{e}^{\mathrm{i}z}}{z^4}\, \mathrm{d}z \right| \leqslant \pi R^{-3}, \quad \text{(iv)} \left| \int_{[0,1+\mathrm{i}]} (z^2+1)^{-1}\, \mathrm{d}z \right| \leqslant \sqrt{2}.$$

10.6 Let $\gamma$ be a square contour such that $\gamma^*$ has its vertices at $(\pm 1 \pm \mathrm{i})R$. Obtain an upper bound for $\left|\int_\gamma z^n\, \mathrm{d}z\right|$ when (i) $n \in \mathbb{Z}$, $n \geqslant 0$ and (ii) $n \in \mathbb{Z}$, $n < 0$.

10.7 Let $p(z)$ be a polynomial. Use the Fundamental theorem of calculus to prove that $\int_\gamma p(z)\, \mathrm{d}z = 0$ for every closed path in $\mathbb{C}$. (This is a very special case of Cauchy's theorem, considered in Chapters 11 and 12.) Deduce that there exists $\varepsilon > 0$ such that, for every polynomial $p(z)$,

$$\left| p(z) - \frac{1}{z} \right| \geqslant \varepsilon \quad \text{whenever } |z| = 1.$$

[Thus the function $1/z$ cannot be uniformly approximated on the unit circle by polynomials.]

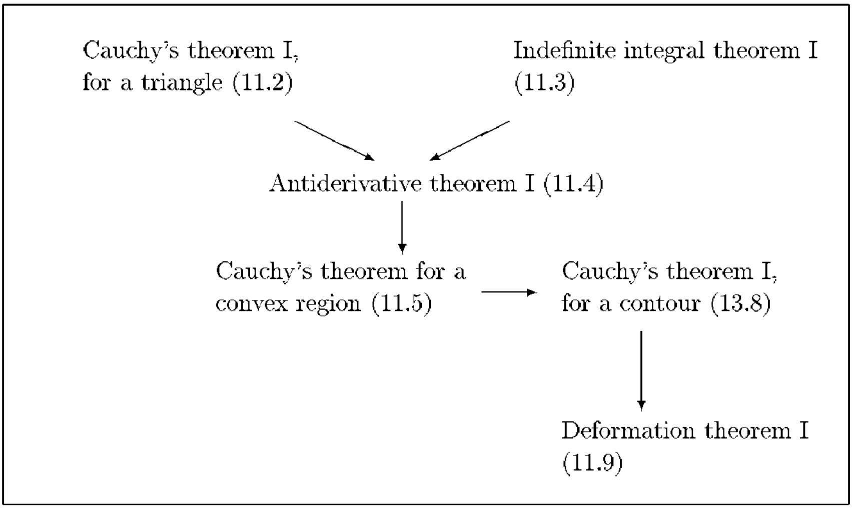
# 11 Cauchy's theorem: basic track

Cauchy's theorem is the centrepiece of complex analysis. It states that

$$\int_\gamma f(z)\,\mathrm{d}z = 0$$

under appropriate conditions on the function $f$, the (closed) path $\gamma$, and the set $G$ on which $f$ is holomorphic. Inextricably bound up with Cauchy's theorem are the Deformation theorem, concerning the equality of $\int_\gamma f(z)\,\mathrm{d}z$ and $\int_{\widetilde{\gamma}} f(z)\,\mathrm{d}z$ under appropriate conditions, and the Antiderivative theorem, dealing with the existence of $F \in \mathrm{H}(G)$ such that $F' = f$. This basic track chapter presents entry-level forms of all three theorems. Table 11.1 indicates the rather convoluted route we take to arrive at these key results.

**Table 11.1**




In this chapter and the next we present Cauchy's theorem in three different forms: Cauchy's theorem I focuses on conditions on $\gamma$, Cauchy's theorem II on conditions on $G$, while Cauchy's theorem III is a definitive, global, version of the theorem. Both theorems I and II are obtained first under restrictive hypotheses. Associated theorems (deformation theorem, for example) are labelled in a corresponding way.

## Cauchy's theorem

**11.1 Contours: a recap.** Recall that in 4.3 we defined a **contour** to be a simple closed path whose image is made up of a finite number of line segments and circular arcs. Our purpose in introducing contours was to avoid the geometrical complexity that general closed paths may exhibit. In particular, we proved that a contour $\gamma$ has an inside $\mathbf{I}(\gamma)$ and an outside $\mathbf{O}(\gamma)$ (Theorem 4.6, a restricted form of the Jordan curve theorem).

As our experience with the Fundamental integral (10.4) indicates, the conclusion of Cauchy's theorem may fail for a function which is not holomorphic *inside* the contour as well as behaving well on it: $z^{-1}$ is holomorphic at every point of the unit circle yet $\int_{\gamma(0;1)} z^{-1}\,\mathrm{d}z = 2\pi\mathrm{i} \neq 0$. We say that $f$ is **holomorphic inside and on** a contour $\gamma$ if $f \in \mathrm{H}(G)$ for some open set $G$ such that $\gamma^* \cup \mathbf{I}(\gamma) \subseteq G$. (Remember our emphasis on *open* sets in the definition of holomorphy in 5.7.)

The Fundamental theorem of calculus (10.8) implies that $\int_\gamma F'(z)\,\mathrm{d}z = 0$ when $\gamma$ is a closed path in an open set $G$, for suitable functions $F$ defined in $G$. A natural way to approach Cauchy's theorem is therefore to find conditions under which $f \in \mathrm{H}(G)$ has an antiderivative $F$ (that is, $f = F'$). It turns out that, provided $G$ is convex, this is so if $\int_\gamma f(z)\,\mathrm{d}z = 0$ for all triangles $\gamma$ in $G$. Consequently we shall first prove that Cauchy's theorem is true in the special case that the path of integration is a triangle.

**11.2 Cauchy's theorem I (for a triangle).** Suppose that $f$ is holomorphic on an open set $G$ which contains a triangle $\gamma$ and the region inside it. Then $\int_\gamma f(z)\,\mathrm{d}z = 0$.

**Proof** We first outline the ideas in the proof. The Fundamental theorem of calculus shows that $\int_{\tilde{\gamma}} p(z)\,\mathrm{d}z = 0$ for any polynomial $p(z)$ and any triangular contour $\tilde{\gamma}$. Near a point $Z$, we can approximate our holomorphic function $f$ by the polynomial $p(z) = f(Z) + (z - Z)f'(Z)$ (see 5.11). Hence we aim to replace

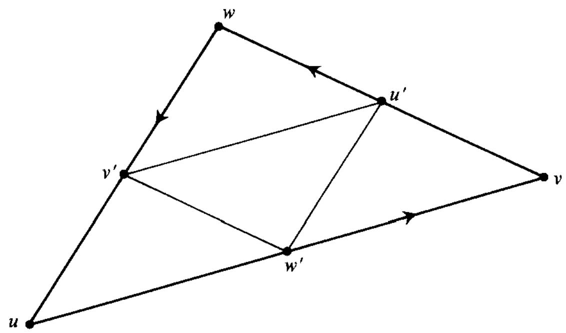



**Figure 11.1** Subdivision of a triangle

$\int_\gamma f(z)\,\mathrm{d}z$ by the sum of integrals round small triangles on the image of each of which $p(z)$ is a good approximation to $f(z)$.

For any distinct points $p$, $q$, $r$, let $[p,q,r]$ denote the triangle formed by joining $[p,q]$, $[q,r]$, and $[r,p]$. Let $\gamma$ be $[u,v,w]$ and $u'$, $v'$, and $w'$ be, respectively, the midpoints of $[v,w]$, $[w,u]$, and $[u,v]$, as shown in Fig. 11.1. Consider the triangles $\gamma^0 = [u',v',w']$, $\gamma^1 = [u,w',v']$, $\gamma^2 = [v,u',w']$, and $\gamma^3 = [w,v',u']$. Then, by 10.5,

$$I := \int_\gamma f(z)\,\mathrm{d}z = \sum_{k=0}^{3} \int_{\gamma^k} f(z)\,\mathrm{d}z.$$

For at least one value of $k$,

$$\left| \int_{\gamma^k} f(z)\,\mathrm{d}z \right| \geqslant \frac{1}{4}\,|I| \quad \text{(by 1.9(3))}.$$

Relabel such a $\gamma^k$ as $\gamma_1$. Repeat the argument with $\gamma_1$ in place of $\gamma$. Proceeding in this way, generate a sequence $\gamma_0, \gamma_1, \gamma_2, \ldots$ of triangles such that

(i) $\gamma_0 = \gamma$,

(ii) for each $n$, $\Delta_{n+1} \subseteq \Delta_n$, where $\Delta_n$ is the closed triangular area having $\gamma_n^*$ as its boundary,

(iii) $\text{length}(\gamma_n) = 2^{-n}L$, where $L = \text{length}(\gamma)$, and

(iv) $4^{-n}\,|I| \leqslant \left|\int_{\gamma_n} f(z)\,\mathrm{d}z\right|$ for all $n \geqslant 0$.

The set $\bigcap_{n=0}^{\infty} \Delta_n$ contains a point $Z$ common to all the triangles $\Delta_n$. (To prove this, select for each $n$ some point $z_n \in \Delta_n$. The sequence $\{z_n\}$ is bounded since all points $z_n$ belong to $\Delta_0$. By Theorem 3.22, $\{z_n\}$ has a subsequence convergent to some point $Z$. For each $n$, $Z$ is a limit point of the subset $\{z_k\}_{k\geqslant n}$ of $\Delta_n$ and so belongs to $\Delta_n$ (see 3.7).)

Fix $\varepsilon > 0$. The function $f$ is differentiable at $Z$, so, for some $r$,

$$|f(z) - f(Z) - (z - Z)f'(Z)| < \varepsilon\,|z - Z| \quad \text{for all } z \in \mathrm{D}(Z; r). \tag{1}$$

Choose $N$ such that $\mathrm{D}(Z;r) \supseteq \Delta_N$. For such $N$,

$$|z - Z| \leqslant 2^{-N}L \quad \text{for all } z \in \Delta_N, \tag{2}$$

by (iii), and

$$\int_{\gamma_N} (f(Z) + (z - Z)f'(Z))\,\mathrm{d}z = 0 \tag{3}$$

by 10.8. Hence, by (1), (2), (3), and 10.10,

$$\left|\int_{\gamma_N} f(z)\,\mathrm{d}z\right| \leqslant \varepsilon(2^{-N})L \times \operatorname{length}(\gamma_N) = \varepsilon(2^{-N}L)^2.$$

By (iv), $|I| \leqslant \varepsilon L^2$. Since $\varepsilon$ is arbitrary, $I = 0$, as required. □

**11.3 Indefinite integral theorem I.** Let $f$ be a continuous complex-valued function on a convex region $G$ such that $\int_\gamma f(z)\,\mathrm{d}z = 0$ for any triangle $\gamma$ in $G$. Let $a$ be an arbitrary fixed point of $G$. Then $F$ defined by

$$F(z) = \int_{[a,z]} f(w)\,\mathrm{d}w$$

is holomorphic in $G$, with $F' = f$.

**Proof** Fix $z \in G$ and let $\mathrm{D}(z;r) \subseteq G$, so that $|h| < r$ implies $z + h \in G$. We show that $(F(z+h) - F(z))/h \to f(z)$ as $h \to 0$. For $|h| < r$, the line segments $[a,z]$, $[z, z+h]$, and $[a, z+h]$ all lie in $G$, since $G$ is convex; see Fig. 10.1. By hypothesis, the integral of $f$ round the triangular contour $[a, z, z+h]$ is zero.

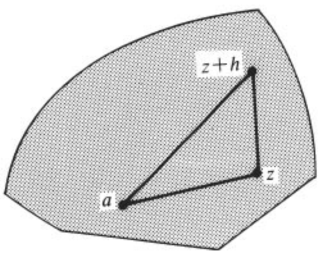



**Figure 11.2** Proof of Indefinite integral theorem I

Hence, by 10.5 and 10.6,

$$F(z+h) - F(z) = \int_{[a,z+h]} f(w)\,\mathrm{d}w - \int_{[a,z]} f(w)\,\mathrm{d}w = \int_{[z,z+h]} f(w)\,\mathrm{d}w.$$

Also, by parametrization, $\int_{[z,z+h]} 1\,\mathrm{d}w = h$. Hence

$$\begin{aligned}\left|\frac{F(z+h)-F(z)}{h} - f(z)\right| &= \frac{1}{|h|}\left|\int_{[z,z+h]} (f(w) - f(z))\,\mathrm{d}w\right| \\ &\leqslant \frac{1}{|h|} \times |h| \times \sup_{w\in[z,z+h]} |f(w) - f(z)| \quad \text{(by 10.10)}\end{aligned}$$

and this tends to zero as $h \to 0$, by continuity of $f$ at $z$. □

**11.4 Antiderivative theorem I.** Let $G$ be a convex region and let $f \in \mathrm{H}(G)$. Then there exists $F \in \mathrm{H}(G)$ such that $F' = f$.

**Proof** Combine Theorems 11.2 and 11.3. □

**11.5 Cauchy's theorem for a convex region.** Let $G$ be a convex region and let $f \in \mathrm{H}(G)$. Then $\int_\gamma f(z)\,\mathrm{d}z = 0$ for every closed path $\gamma$ in $G$.

**Proof** Combine Theorem 11.4 and the Fundamental theorem of calculus. □

Often the region in which we want to apply Cauchy's theorem is not convex. For example, this is the case when we use keyhole-shaped contours in cut planes (see Fig. 19.3). We now present the version of Cauchy's theorem we use when Theorem 11.5 is not applicable.

**11.6 Cauchy's theorem I (for a contour).** Suppose that $f$ is holomorphic inside and on a closed contour $\gamma$. Then $\int_\gamma f(z)\,\mathrm{d}z = 0$.

**Proof** Suppose first that $\gamma$ is a polygon. By triangulating $\gamma$ (see 4.8) we can write $\int_\gamma f(z)\,\mathrm{d}z$ as $\sum_{k=1}^{N} \int_{\gamma_k} f(z)\,\mathrm{d}z$, where each $\gamma_k$ is a triangle; note that the integrals along the inserted line segments cancel. By Theorem 11.2, the integral of $f$ along each $\gamma_k$ is zero, so $\int_\gamma f(z)\,\mathrm{d}z = 0$.

Now let $\gamma$ be any contour, and let $G$ be an open set containing $\gamma^* \cup \mathbf{I}(\gamma)$ on which $f$ is holomorphic. We shall 'approximate' $\gamma$ by a polygonal contour. To do this we cover $\gamma^*$ with overlapping discs $D_k = \mathrm{D}(\gamma(t_k); m)$ ($k = 0, \ldots, N; t_0 < t_1 < \cdots < t_N$; $\gamma(t_0) = \gamma(t_N)$) which satisfy the conditions (i)–(iv) of the Covering theorem (4.5). By increasing the number of discs if necessary we may

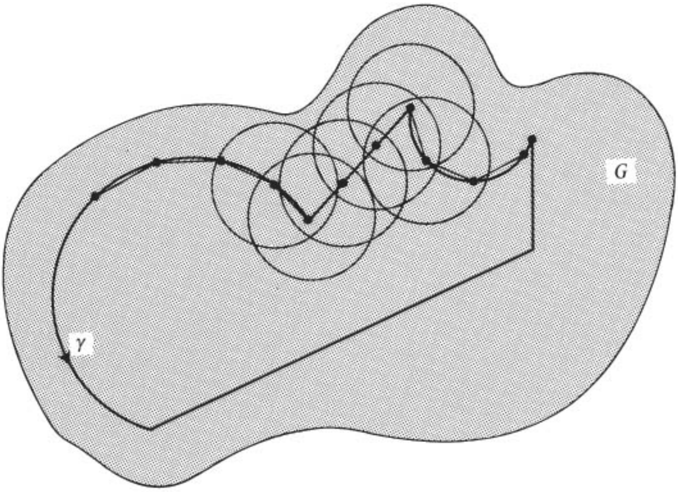



**Figure 11.3** Proof of Cauchy's theorem I

assume that each $\gamma_k := \gamma\lceil[t_k, t_{k+1}]$ is a line segment or a circular arc, and also that the line segments $\widetilde{\gamma}_k := [\gamma(t_k), \gamma(t_{k+1})]$ $(k = 0, \ldots, N-1)$ join to form a polygonal contour $\widetilde{\gamma}$ such that $\widetilde{\gamma}^* \cup \mathbf{I}(\widetilde{\gamma})$ is contained in $\bigcup_{k=0}^{N} D_k \cup \mathbf{I}(\gamma)$, and so in $G$; see Fig. 11.3. We have $\int_{\widetilde{\gamma}} f(z)\,\mathrm{d}z = 0$. Furthermore, for each $k$, the join of $\gamma_k$ and $-\widetilde{\gamma}_k$ is a closed path in $D_k$, which is convex. By 11.5 and 10.5,

$$\int_{\gamma_k} f(z)\,\mathrm{d}z = \int_{\widetilde{\gamma}_k} f(z)\,\mathrm{d}z \quad \text{for each } k.$$

Hence

$$\int_{\gamma} f(z)\,\mathrm{d}z = \sum_{k=0}^{N-1} \int_{\gamma_k} f(z)\,\mathrm{d}z = \sum_{k=0}^{N-1} \int_{\widetilde{\gamma}_k} f(z)\,\mathrm{d}z = \int_{\widetilde{\gamma}} f(z)\,\mathrm{d}z = 0. \qquad \square$$

**11.7 Evaluation of integrals: taking stock.** We began in Chapter 10 by evaluating integrals from scratch by recourse to the parametric definition given in 10.3. This entails splitting the integrand into real and imaginary parts, and then using the real Fundamental theorem of calculus. This may lead to laborious manipulations or, at worst, to integrals we cannot evaluate.

We graduated to the complex Fundamental theorem of calculus (10.8). To apply this we must recognize our integrand as a continuous derivative, and we cannot always do this.

With Cauchy's theorem available we see instantly that our hard-won answer in Example 10.7 is right. We also now know, for example, that $\int_\gamma e^{z^2}\,dz = 0$ for any closed contour $\gamma$. You would not be able to obtain this result without Cauchy's theorem! There are situations where Cauchy's theorem is not applicable but where either parametrization or the Fundamental theorem of calculus can be applied.

In Chapters 13 and 18 we shall greatly extend our range of techniques for evaluating complex integrals. These techniques have their roots in the results of this chapter.

**11.8 Examples (to illustrate 11.7).** We claim that $I := \int_{\gamma(0;1)} f(z)\,dz = 0$ for each of the functions $f$ below.

- Take $f(z) = 1/z^2$. Here $I = 0$ by 10.4.
- Take $f(z) = \operatorname{cosec}^2 z$. Then $f(z) = (d/dz)\cot z$ in an open set containing $\gamma(0;1)^*$ so $I = 0$ by the Fundamental theorem of calculus. Cauchy's theorem is not applicable because $\sin z = 0$ at $0 \in \mathbf{I}(\gamma(0;1))$.
- Take $f(z) = (z^2+4)^{-1}e^{iz^2}$. In this case, the zeros of the denominator do not lie inside or on $\gamma(0;1)$. Hence $I = 0$ by Cauchy's theorem.
- Take $f(z) = (\operatorname{Im} z)^2$. Here we are dealing with a nowhere holomorphic function (the Cauchy–Riemann equations fail except at 0; remember 5.14). The only technique available is parametrization. Write $\gamma(t) = e^{it}$ $(0 \leqslant t \leqslant 2\pi)$. By de Moivre's theorem, $f(\gamma(t)) = \sin 2t$. So

$$I = \int_0^{2\pi} (\sin^2 t)ie^{it}\,dt = \int_0^{2\pi} -2\cos t\sin^2 t\,dt + 2i\int_0^{2\pi} \sin t\cos^2 t\,dt = 0.$$

## Deformation

We next consider deformation of paths at a basic track level. The objective is to be able to replace the integral of a given function $f$ along some given path by the integral of the same function along a more amenable path. Here we present forms of the Deformation theorem sufficient for most needs. [On the advanced track, deformation is approached via the topological concept of homotopy (see 12.1).]

The hypotheses in 11.9(1) allow us to replace an integral round a contour $\gamma$ by an integral round a circle, centre $a$, inside $\gamma$. We use this result in the proofs of important theorems in Chapter 13. In these applications, the integrand involves a factor of $(z-a)^{-n}$ $(n = 1, 2, \ldots)$; this explains why we must permit non-holomorphy at one point $a$ inside our contour. Another commonly needed type of deformation is that in 11.9(3). Loosely, this version of the theorem gives sufficient conditions for the integral of a holomorphic function $f$ along a circline path with fixed endpoints to be independent of the path chosen.

**11.9 Deformation theorem I.**

(1) Suppose that $\gamma$ is a positively oriented contour and that $\overline{\mathrm{D}}(a;r) \subseteq \mathbf{I}(\gamma)$. Let $f$ be holomorphic inside and on $\gamma$ except possibly at $a$. Then

$$\int_\gamma f(z)\,\mathrm{d}z = \int_{\gamma(a;r)} f(z)\,\mathrm{d}z.$$

(2) Suppose that $\gamma$ and $\widehat{\gamma}$ are positively oriented contours such that $\widehat{\gamma}$ lies inside $\gamma$, that is, $\widehat{\gamma}^* \cup \mathbf{I}(\widehat{\gamma}) \subseteq \mathbf{I}(\gamma)$. Let $f$ be holomorphic inside and on $\gamma$. Then $\int_\gamma f(z)\,\mathrm{d}z = \int_{\widehat{\gamma}} f(z)\,\mathrm{d}z$.

(3) Suppose that $\gamma_1$ and $\gamma_2$ are circline paths with a common initial point and common final point, let $\gamma := \gamma_1 \cup (-\gamma_2)$, and suppose that $\gamma$ is simple. Let $f$ be holomorphic inside and on $\gamma$. Then $\int_{\gamma_1} f(z)\,\mathrm{d}z = \int_{\gamma_2} f(z)\,\mathrm{d}z$.

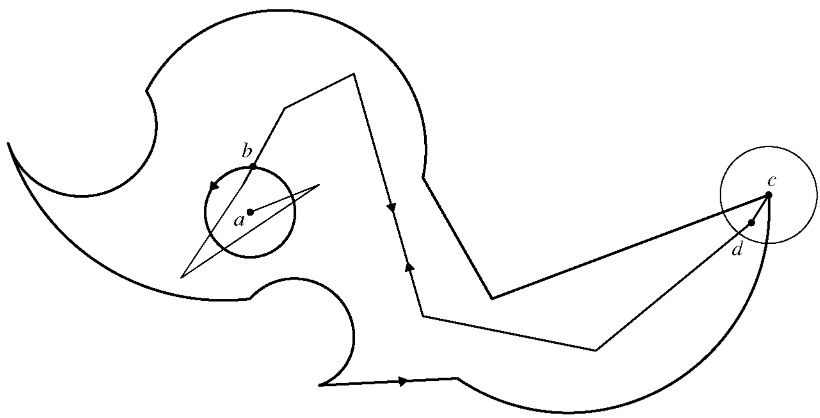



**Figure 11.4** The proof of Deformation theorem I

**Proof** (1) The strategy is to form a closed path to which Cauchy's theorem can be applied. Let $c$ be the initial point of $\gamma$ and let $\mathrm{D}(c;\delta)$ be such that $f \in \mathrm{H}(\mathrm{D}(c;\delta))$. Then $\mathbf{I}(\gamma) \cap \mathrm{D}(c;\delta) \neq \varnothing$. Take $d \in \mathbf{I}(\gamma) \cap \mathrm{D}(c;\delta)$. Since $\mathbf{I}(\gamma)$ is polygonally connected (Theorem 4.6) there is a polygonal path $\gamma_1$ in $\mathbf{I}(\gamma)$ joining $d$ to $a$ and we can assume that this is simple. Let the parameter interval of $\gamma_1$ be $[\alpha,\beta]$. There is a point $b = \gamma_1(T)$ on $\gamma_1^*$ such that $|\gamma_1(t) - a| > r$ for $0 \leqslant t < T$ (so that $b$ is the first point at which $\gamma_1^*$ meets the circle $|z-a| = r$. Then $\gamma_2 := [c,d] \cup \gamma_1$ joins $c$ to $b$. Now define $\Gamma$ to be the join of $\gamma$, $\gamma_2$, $-\gamma(a;r)$, and $-\gamma_2$. This fails to be a contour solely because of the double-tracing of $\gamma_2^*$. However, it is easy to see that the proof of Theorem 11.6 remains valid for a path such as $\Gamma$ so that $\int_\Gamma f(z)\,\mathrm{d}z = 0$. Finally, by 10.5 and 10.6, we obtain $\int_\gamma f(z)\,\mathrm{d}z = \int_{\gamma(a;r)} f(z)\,\mathrm{d}z$.

We obtain (2) from (1) by choosing some disc $\mathrm{D}(a;r)$ within $\mathbf{I}(\widehat{\gamma})$ and applying (1) twice to obtain $\int_\gamma f(z)\,\mathrm{d}z = \int_{\gamma(a;r)} f(z)\,\mathrm{d}z = \int_{\widehat{\gamma}} f(z)\,\mathrm{d}z$.

For (3) we can apply Cauchy's theorem 11.6, together with 10.5(1) and 10.6. □

**11.10 The Fundamental integral revisited.** In 10.4 we computed $\int_\gamma (z-a)^n\,\mathrm{d}z$ for $n \in \mathbb{Z}$ and $\gamma = \gamma(a;r)$. We can now evaluate this integral for other choices of $\gamma$. Assume that $\gamma$ is a positively oriented contour and that $a \notin \gamma^*$. Then

$$\int_\gamma (z-a)^n\,\mathrm{d}z = \begin{cases} 0 & \text{if } a \in \mathbf{O}(\gamma), \\ 0 & \text{if } a \in \mathbf{I}(\gamma) \text{ and } n \neq -1, \\ 2\pi\mathrm{i} & \text{if } a \in \mathbf{I}(\gamma) \text{ and } n = -1. \end{cases}$$

For $n \neq -1$ this comes from the Fundamental theorem of calculus (10.8). For $n = -1$ we do not have an antiderivative for $(z-a)^n$ so cannot use the Fundamental theorem of calculus. However Cauchy's theorem I is applicable when $a \in \mathbf{O}(\gamma)$ and we can use 10.4 together with Deformation theorem I (11.9(2)) when $a \in \mathbf{I}(\gamma)$.

**11.11 Example (Deformation theorem).**

- Consider $I = \int_{\gamma(0;2)} f(z)\,\mathrm{d}z$ when $f(z) = 2(4z^2-1)^{-1}$. Write $f(z)$ as $(2z-1)^{-1} - (2z+1)^{-1}$. We can now invoke the Deformation theorem (11.9(2)) and then 10.4:

$$I = \int_{\gamma(1/2;1/4)} \frac{1}{2(z-\frac{1}{2})}\,\mathrm{d}z - \int_{\gamma(-1/2;1/4)} \frac{1}{2(z+\frac{1}{2})}\,\mathrm{d}z = \frac{1}{2}.2\pi\mathrm{i} - \frac{1}{2}.2\pi\mathrm{i} = 0.$$

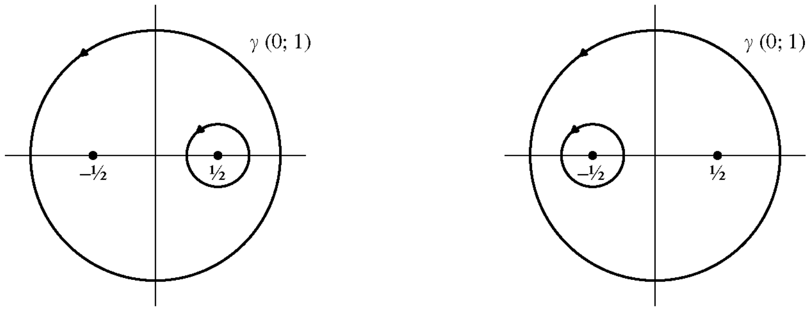



**Figure 11.5** Illustrating Example 11.11

## Logarithms again

We showed in 7.14 that for each $z \neq 0$ we can find infinitely many solutions to the equation $e^w = z$, differing by integer multiples of $2\pi i$. We have seen that judicious selection of values produces a well-defined logarithm function, provided we restrict the argument of $z$ suitably (see 7.14). We now have enough machinery to investigate logarithms in more depth.

**11.12 Multifunctions from indefinite integrals.** It is well known that the real logarithm is given by

$$\log x = \int_1^x \frac{1}{s}\, ds \quad (x \in (0, \infty)).$$

Is there an analogous description for the complex logarithm? And how might the many-valuedness manifest itself?

Let $0 \neq z = |z|\, e^{i\theta}$, with $-\pi < \theta \leqslant \pi$, so that the value of $\theta$ is uniquely determined by $z$. Essentially, we are working in the plane with a cut along $(-\infty, 0]$, using 'upper-edge' values on the cut. Let

$$F_0(z) := \int_{\Gamma(z)} \frac{1}{w}\, dw,$$

where the path $\Gamma(z)$ is the join of $\Gamma_1 := [1, |z|]$ and $\Gamma_2$ defined by

$$\Gamma_2(z) := \begin{cases} |z|\, e^{it} & (t \in [0, \theta]), \quad \text{if } \operatorname{Im} z \geqslant 0, \\ |z|\, e^{i(\theta - t)} & (t \in [\theta, 0]), \quad \text{if } \operatorname{Im} z < 0; \end{cases}$$

see Fig. 11.6. Note that the path does not cross the cut.

Suppose that $\operatorname{Im} z \geqslant 0$ and write $z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$. Then

$$F_0(z) := \int_{\Gamma(z)} \frac{1}{w}\,\mathrm{d}w = \int_0^{|z|} \frac{1}{s}\,\mathrm{d}s + \int_0^{\theta} \frac{1}{|z|\,\mathrm{e}^{\mathrm{i}t}}\,|z|\,\mathrm{i}\mathrm{e}^{\mathrm{i}t}\,\mathrm{d}t = \log|z| + \mathrm{i}\theta.$$

Similarly, we can show that $F_0(z) = \log|z| + \mathrm{i}\theta$ if $\operatorname{Im} z < 0$.

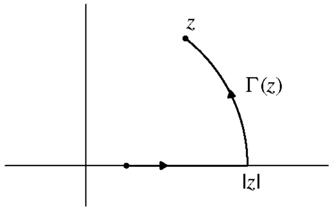



**Figure 11.6** The path $\Gamma(z)$

So we have obtained a valid integral formula for the complex logarithm. To achieve this we chose a particular circline path $\Gamma(z)$ from 1 to $z$. What if we had chosen a different circline path in the cut plane, or a path in $\mathbb{C} \setminus \{0\}$? Recall that for any positively oriented contour $\gamma$,

$$\int_\gamma \frac{1}{w}\,\mathrm{d}w = \begin{cases} 0 & \text{if } w \in \mathbf{O}(\gamma), \\ 2\pi\mathrm{i} & \text{if } w \in \mathbf{I}(\gamma) \end{cases}$$

(that Fundamental integral again!). For $k \in \mathbb{Z}$, let

$$\Gamma_k(z) := \begin{cases} \overbrace{\left(\gamma(0;1) \cup \cdots \cup \gamma(0;1)\right)}^{k \text{ times}} \cup \Gamma(z) & \text{if } k \geqslant 0, \\ \overbrace{\left((-\gamma(0;1)) \cup \cdots \cup (-\gamma(0;1))\right)}^{-k \text{ times}} \cup \Gamma(z) & \text{if } k < 0. \end{cases}$$

Then, for $z \neq 0$,

$$\int_{\Gamma_k(z)} \frac{1}{w}\,\mathrm{d}w = F_0(z) + 2k\pi\mathrm{i}.$$

If we were to replace $\gamma(0;1)$ (which encircles 0) by, for example, $\gamma(1/2;1/2)$ (which does not encircle 0), then the factor $2k\pi\mathrm{i}$ would not appear. This can be generalized: suppose we allow $\gamma(z)$ to be any circline path from 1 to $z$ in $\mathbb{C} \setminus \{0\}$. Then

$$\int_{\gamma(z)} \frac{1}{w}\,\mathrm{d}w$$

takes precisely the values in $[\![\log z]\!]$, as $\gamma(z)$ varies.

Other multifunctions may also be viewed in terms of indefinite integrals. For example, the inverse tangent $[\![\tan^{-1} z]\!]$ is given by the integral

$$\int_{\gamma(z)} \frac{1}{1+w^2}\,\mathrm{d}w.$$

Here $\gamma(z)$ is any path from 0 to $z$ not passing through $\pm\mathrm{i}$ (branch points). The multiple values arise from circuits round i or $-\mathrm{i}$ or both. Put cuts along the imaginary axis as indicated in Fig 11.7. In the cut plane, the indefinite integral defines a holomorphic branch of $\tan^{-1}$. Other inverse trigonometric and hyperbolic functions can be treated in a similar way.

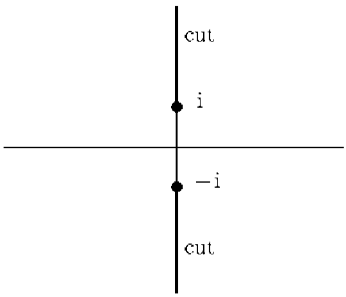



**Figure 11.7** Cut plane for the inverse tangent

The discussion in 11.12 indicates that it is encirclement of 0 that must be prevented if we are to have a well-defined logarithm function. Once we have done this, the logarithm locally behaves well, as the next theorem shows. A more general, and definitive, version of this result is given in 12.7.

**11.13 Theorem (logarithm in a convex region).** Suppose that $G$ is a convex region not containing 0. Then there exists a function $f = \log_G \in \mathrm{H}(G)$ such that $\mathrm{e}^{f(z)} = z$ for all $z \in G$ and

$$f(z) - f(a) = \int_\gamma \frac{1}{w}\,\mathrm{d}w \quad \text{for all } a \text{ and } z \text{ in } G,$$

where $\gamma$ is any path in $G$ with endpoints $a$ and $z$. The function $f$ is uniquely determined up to the addition of an integer multiple of $2\pi\mathrm{i}$. Furthermore, for each $z \in G$,

$$\log_G z = \log|z| + \mathrm{i}\theta(z),$$

where $\theta(z) \in [\![\arg z]\!]$ and $z \mapsto \theta(z)$ is a continuous function in $G$.

**Proof** By the Antiderivative theorem I (11.4) there exists $f \in \mathrm{H}(G)$ such that $f'(z) = 1/z$ for all $z \in G$. Using the Chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}z} z\mathrm{e}^{-f(z)} = \mathrm{e}^{-f(z)} - zf'(z)\mathrm{e}^{-f(z)} = 0.$$

Hence, by 5.12(1), $z = C\mathrm{e}^{f(z)}$, where $C$ is a non-zero constant. By adding a suitable constant to $f$, we may assume that $C = 1$. The integral formula for $f$ comes from the proof of the Antiderivative theorem. Suppose that, for all $z \in G$, we have $\mathrm{e}^{f(z)} = \mathrm{e}^{g(z)}$, where both $f$ and $g$ are holomorphic in $G$. Then $f - g$ has zero derivative and so, by 5.12(1) again, is a constant $K$. We have $\mathrm{e}^K = 1$, so $K = 2n\pi\mathrm{i}$ for some integer $n$ (see 7.10).

The last part comes from 7.14 and the fact that the imaginary part of a holomorphic function must be continuous. □

## Exercises

11.1 Each of the following integrals is zero:

$$\text{(i)} \int_{\gamma(1;1)} \frac{1}{z-3}\,\mathrm{d}z, \qquad \text{(ii)} \int_{\gamma(\mathrm{i};4)} \frac{1}{(z-3)^3}\,\mathrm{d}z,$$
$$\text{(iii)} \int_{\gamma(0;1)} z\,|z|^4\,\mathrm{d}z, \qquad \text{(iv)} \int_{\gamma(1;1)} (1+\mathrm{e}^z)^{-1}\,\mathrm{d}z.$$

Give a reason (or reasons) in each case.

11.2 (a) Let $\gamma$ be $\gamma(0;2)$. For which of the following functions $f$ is $\int_\gamma f(z)\,\mathrm{d}z$ equal to zero?

$$\text{(i) } \overline{z}, \quad \text{(ii) } \frac{1}{z-1}, \quad \text{(iii) } z^5 \sin^3 z, \quad \text{(iv) } \sec^2 z.$$

Give reasons.

(b) Repeat (a) with $\gamma$ the contour described in Exercise 4.1(iv).

11.3 Evaluate $\int_\gamma (1+z^2)^{-1}\,\mathrm{d}z$ when $\gamma$ is

$$\text{(i) } \gamma(1;1), \quad \text{(ii) } \gamma(\mathrm{i};1), \quad \text{(iii) } \gamma(-\mathrm{i};1), \quad \text{(iv) } \gamma(0;2), \quad \text{(v) } \gamma(3\mathrm{i};\pi).$$

(Hint: this is an exercise on Cauchy's theorem and the Deformation theorem together with 10.4; partial fractions will be helpful for some of the parts.)

11.4 Define a path $\gamma$ whose image $\gamma^*$ is the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

traced anticlockwise. By showing that $\int_\gamma f(z)\,\mathrm{d}z = \int_{\widehat{\gamma}} f(z)\,\mathrm{d}z$ for a suitable circle $\widehat{\gamma}$, prove that

$$\int_0^{2\pi} \frac{1}{a^2\cos^2 t + b^2 \sin^2 t}\,\mathrm{d}t = \frac{2\pi}{ab} \quad (a > 0,\ b > 0).$$

(Hint: this would be an immediate consequence of the Deformation theorem I (11.9(2)) if we had not restricted attention there to circline paths; prove with the aid of 11.5 that deformation of the elliptical contour to a circle is permissible.)

11.5 A subset $S$ of $\mathbb{C}$ is said to be **star-shaped** if there exists $a \in S$ such that $[a, z] \subseteq S$ for all $z \in S$.

(i) Prove that a convex set is star-shaped, and exhibit a non-convex set which is star-shaped.

(ii) Prove that an open star-shaped set is a region.

(iii) Let $G$ be open and star-shaped and let $f \in \mathrm{H}(G)$. Adapt the proof of 11.3 to prove that $\int_\gamma f(z)\,\mathrm{d}z = 0$ (that is, prove Cauchy's theorem for a star-shaped region).

11.6 Establish which of the following sets are star-shaped:

(i) $\{ z : \operatorname{Im} z > 0 \}$,
(ii) $\{ z : 1 < |z| < 2 \}$,
(iii) $\{ z : |z-2| > 3,\ |z| < 2 \}$,
(iv) $\mathbb{C}_\pi = \mathbb{C} \setminus (-\infty, 0]$,
(v) $\mathbb{C} \setminus \{\pm 1\}$,
(vi) $S_{\alpha,\beta}$ $(\beta - \alpha < 2\pi)$ (as in 2.6),
(vii) $\mathbb{C} \setminus \{ z : |z| = 1,\ \operatorname{Re} z \geqslant 0 \}$.

11.7 Let $\gamma$ be a circline path with initial point 0 and final point 1. Find all the possible values of

$$\text{(i)} \int_\gamma z^3\,\mathrm{d}z, \quad \text{(ii)} \int_\gamma (1+z^2)^{-1}\,\mathrm{d}z \text{ (where } \pm 1 \notin \gamma^*).$$

# 12 Cauchy's theorem: advanced track

We approached Cauchy's theorem in a topologically unsophisticated way in Chapter 11. The various versions of the theorem we have presented so far are adequate for applications, but none of these is a definitive formulation. To treat Cauchy's theorem only in a utilitarian way would be demeaning to complex analysis. So we now delve a little deeper, hinting at the notions which lie at the heart of Cauchy's theorem.

## Deformation and homotopy

We introduced polygonal connectedness as a means of distinguishing topologically between a single open disc and the disjoint union of more than one open disc. We now seek to employ paths to distinguish an open disc $D$ from an open annulus $A$. Informally, of course, the essential difference is that $A$ has a hole in it but $D$ does not. This means that a closed path in $A$ whose image encircles the hole cannot be shrunk, within $A$, to a point, while every closed path in $D$ can be deformed to a point. We now make precise what we mean by deformation of closed paths.

**12.1 Deformation.** Let $G$ be a non-empty open set in $\mathbb{C}$ and let $\gamma$ and $\tilde{\gamma}$ be closed paths in $G$. We say that $\tilde{\gamma}$ can be obtained from $\gamma$ by an **elementary deformation** if there exist open convex subsets $G_0, G_1, \ldots, G_{N-1}$ of $G$ such that $\gamma$ can be expressed as the join of paths $\gamma_0, \gamma_1, \ldots, \gamma_{N-1}$ and $\tilde{\gamma}$ as the join of paths $\tilde{\gamma}_0, \tilde{\gamma}_1, \ldots, \tilde{\gamma}_{N-1}$ in such a way that, for $k = 0, \ldots, N-1$, $\gamma_k$ and $\tilde{\gamma}_k$ lie in $G_k$ and have common initial and final points (see Fig. 12.1).

Furthermore, closed paths $\gamma$ and $\tilde{\gamma}$ in $G$ are said to be **homotopic** (in $G$) if $\tilde{\gamma}$ can be obtained from $\gamma$ via a finite number of elementary deformations.

Elementary deformation is a more natural concept than it might seem from its somewhat fearsome definition. The idea is to cover the images of the two paths by overlapping convex regions and, within each of these, to replace a portion of $\gamma^*$ by a portion of $\tilde{\gamma}^*$. The reason for working with convex sets is that we have a Cauchy theorem for convex regions.

Topologists have a definition of (closed path) homotopy, based on a continuous deformation process, which is, non-trivially, equivalent to our homotopy definition. The underlying idea is to take $G$, $\gamma$, and $\widetilde{\gamma}$ as above and to think of a rubber band positioned over $\gamma^*$. The path $\widetilde{\gamma}$ is homotopic to $\gamma$ if the rubber band can be slid and stretched so as to coincide with $\widetilde{\gamma}^*$ (correctly oriented) without ever moving outside $G$.

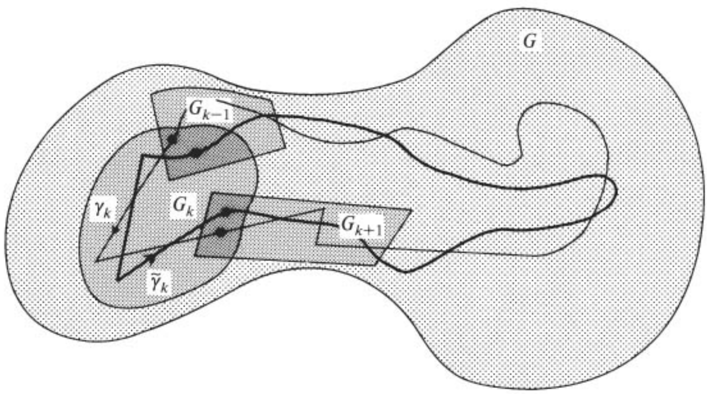



**Figure 12.1** Elementary deformation

**12.2 Definitions (null path, simply connected region).** A path $\gamma$ lying in a set $G$ is said to be **null** if $\gamma^* = \{a\}$ for some $a \in G$. A region $G$ is **simply connected** if every closed path in $G$ is homotopic to a null path in $G$.

**12.3 Examples (simple connectedness).**

- The definition of elementary deformation implies that any two closed paths in a convex region are homotopic. It follows that any convex region is simply connected. In particular, any disc $\mathrm{D}(a;r)$ is simply connected.
- No open annulus is simply connected. This is eminently plausible, but far from elementary to prove. Jumping ahead, Cauchy's theorem II (12.5) tells us that $\int_\gamma f(z)\,\mathrm{d}z = 0$ whenever $f$ is holomorphic in a simply connected region $G$ and $\gamma$ is a closed path in $G$. Take $G = \{\, z : R < |z-a| < S \,\}$ $(0 \leqslant R < S \leqslant \infty)$, $f(z) = (z-a)^{-1}$, and $\gamma = \gamma(a;r)$ $(R < r < S)$. Then $\int_\gamma f(z)\,\mathrm{d}z = 2\pi\mathrm{i} \neq 0$—the Fundamental integral once again.

- For any real number $\alpha$, let

$$\mathbb{C}_\alpha := \{\mathbb{C} \setminus \{ z \in \mathbb{C} \, : \, z = |z| \, e^{i\alpha} \}$$

(so that $\mathbb{C}_\alpha$ is the plane with a ray from 0 excluded). Exercise 12.3 indicates how to prove that $\mathbb{C}_\alpha$ is simply connected.

**12.4 Deformation theorem II.** Suppose that $f$ is holomorphic in an open set $G$ and that $\gamma$ and $\widetilde{\gamma}$ are homotopic closed paths in $G$. Then

$$\int_\gamma f(z)\,dz = \int_{\widetilde{\gamma}} f(z)\,dz.$$

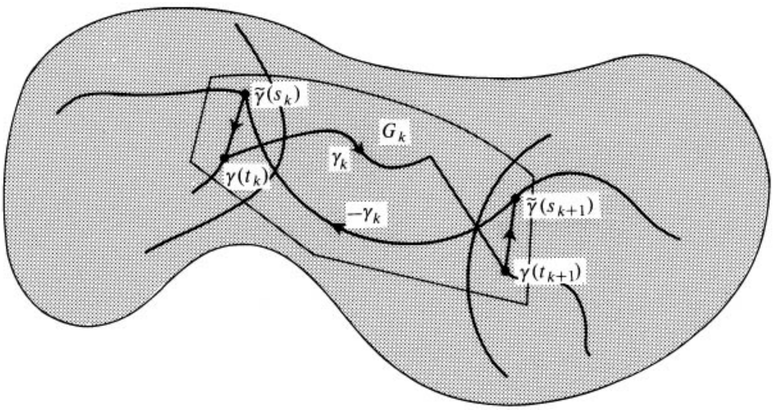



**Figure 12.2** Illustration of the proof of Deformation theorem II

**Proof** We may assume that $\widetilde{\gamma}$ is obtained from $\gamma$ by an elementary deformation. We adopt the notation of 12.1. For each $k = 0, \ldots, N-1$, the join $\Gamma_k$ of $\gamma_k$, $[\gamma(t_{k+1}), \widetilde{\gamma}(s_{k+1})]$, $-\widetilde{\gamma}_k$, and $[\widetilde{\gamma}(s_k), \gamma(t_k)]$ is a closed path in the convex region $G_k$, so Theorem 11.5 implies that $\int_{\Gamma_k} f(z)\,dz = 0$. But

$$\begin{aligned}\int_\gamma f(z)\,dz - \int_{\widetilde{\gamma}} f(z)\,dz &= \sum_{k=0}^{N-1} \left( \int_{\gamma_k} f(z)\,dz - \int_{\widetilde{\gamma}_k} f(z)\,dz \right) \\ &= \sum_{k=0}^{N-1} \int_{\Gamma_k} f(z)\,dz\end{aligned}$$

since the integrals along the line segments cancel. □

## Holomorphic functions in simply connected regions

**12.5 Cauchy's theorem II.** Suppose that $f$ is holomorphic in a simply connected region $G$. Then $\int_\gamma f(z)\,\mathrm{d}z = 0$ for every closed path $\gamma$ in $G$.

**Proof** The simple connectedness of $G$ implies that $\gamma$ is homotopic to a null path $\widehat{\gamma}$ in $G$ (see 12.2). By Deformation theorem II,

$$\int_\gamma f(z)\,\mathrm{d}z = \int_{\widehat{\gamma}} f(z)\,\mathrm{d}z.$$

The latter integral is clearly zero. □

**12.6 Antiderivative theorem II.** Let $G$ be a simply connected region and let $f \in \mathrm{H}(G)$. Then there exists $F \in \mathrm{H}(G)$ such that $F' = f$.

**Proof** We cannot define $F(z) = \int_{[a,z]} f(w)\,\mathrm{d}w$ as we did in the proof of Theorem 11.3, since there may not exist a universal point $a \in G$ such that $[a,z] \subseteq G$ for every $z \in G$ (that is, $G$ need not be star-shaped; see Exercise 11.5). The remedy is to substitute for $[a,z]$ some polygonal path $\gamma(z)$ in $G$ joining a fixed point $a$ to $z$; this is possible by Theorem 3.13. Then, if $\mathrm{D}(a;r) \subseteq G$ and $|h| < r$,

$$\begin{aligned} F(z+h) - F(z) &= \int_{\gamma(z+h)} f(w)\,\mathrm{d}w - \int_{\gamma(z)} f(w)\,\mathrm{d}w \\ &= \int_{[z,z+h]} f(w)\,\mathrm{d}w, \end{aligned}$$

by Lemma 10.5 and Cauchy's theorem II. The proof is completed in the same way as that of Theorem 11.4. □

**12.7 Theorem (logarithm in a simply connected region).** The statements in Theorem 11.13 remain valid for any simply connected region $G$ with $0 \notin G$.

**Proof** We proceed exactly as in 11.13, but appeal to Antiderivative theorem II instead of to Antiderivative theorem I. □

## Argument and index

**12.8 Index.** Let $\gamma$ be a closed path and let $w \notin \gamma^*$. Define the **index** (also called **winding number**) $\mathrm{n}(\gamma, w)$ of $\gamma$ about $w$ by

$$\mathrm{n}(\gamma, w) := \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{1}{z-w} \, \mathrm{d}z.$$

We can restate in terms of index the results concerning the Fundamental integral in the case $n = -1$ that we recorded in 11.10: if $\gamma$ is a positively oriented contour, then $\mathrm{n}(\gamma, w) = 1$ if $w$ lies inside $\gamma$ and 0 if $w$ lies outside $\gamma$. This suggests that we should replace our interim definition of orientation in 4.3 by the following: a contour (or more generally a simple closed path) is **positively oriented** if $\mathrm{n}(\gamma, w) = +1$ for any $w \in \mathbf{I}(\gamma)$. For any closed path $\gamma$, we clearly have $\mathrm{n}(-\gamma, w) = -\mathrm{n}(\gamma, w)$, by 10.5(1).

A contour is, of course, a very special type of closed path: because it is simple, it cannot wind round the same point more than once. The intention is that, for a closed path $\gamma$, the index $\mathrm{n}(\gamma, w)$ should measure the number of times $\gamma$ winds round $w$, taking orientation into account. The next theorem shows that this is a valid interpretation. For $w \notin \gamma^*$, let $\gamma_w(t) = \gamma(t) - w$. Then $\mathrm{n}(\gamma, w) = \mathrm{n}(\gamma_w, 0)$. We may therefore, without loss of generality, take $w = 0$ in what follows.

**12.9 Theorem (properties of index).** Let $\gamma$ be a closed path with parameter interval $[\alpha, \beta]$ and let $0 \notin \gamma^*$. Then

(1) $\mathrm{n}(\gamma, 0)$ is an integer, where $2\pi \mathrm{i}\, \mathrm{n}(\gamma, 0) = \int_\gamma z^{-1} \, \mathrm{d}z$.

(2) There exists a continuous function $\eta \colon [\alpha, \beta] \to \mathbb{R}$, unique up to an integer multiple of $2\pi$, such that
  - (i) $2\pi \mathrm{n}(\gamma, 0) = \eta(\beta) - \eta(\alpha)$;
  - (ii) $\eta(t) \in [\![ \arg(\gamma(t)) ]\!]$ for all $t \in [\alpha, \beta]$.

**Proof** Let $G$ be an open set containing $\gamma^*$, with $0 \notin G$. We should like to use Theorem 12.7, but are prevented from doing so directly because we do not know that $G$ can be chosen to be a simply connected region. We therefore work locally. We construct points $\alpha = t_0 < t_1 < \cdots < t_N = \beta$ and discs $D_0, D_1, \ldots, D_N$ as in the Covering theorem (4.5). For $k = 0, \ldots, N$, let $g_k$ be a holomorphic logarithm in $D_k$, by taking $G = D_k$ in Theorem 11.13 (the general form 12.7 is not needed here). Then $g_k(z) = \log|z| + \mathrm{i}\theta_k(z)$ for $z \in D_k$. For $z \in D_k \cap D_{k-1}$, we have $\theta_{k-1}(z) - \theta_k(z) = 2\pi \mathrm{i} n_k$, where $n_k \in \mathbb{Z}$. Let $z_k = \gamma(t_k)$ $(k = 0, \ldots, N)$. Note that $z_0 = \gamma(\alpha) = \gamma(\beta) = z_N$. Then

$$\begin{aligned}
\mathrm{n}(\gamma, 0) &= \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{1}{z}\,\mathrm{d}z \\
&= \frac{1}{2\pi \mathrm{i}} \sum_{k=0}^{N-1} \int_{[z_k, z_{k+1}]} \frac{1}{z}\,\mathrm{d}z && \text{(by 10.5(2) \& 11.4)} \\
&= \frac{1}{2\pi \mathrm{i}} \sum_{k=0}^{N-1} \big(g_k(z_{k+1}) - g_k(z_k)\big) && \text{(by 12.7)} \\
&= \frac{1}{2\pi} \sum_{k=0}^{N-1} \big(\theta_k(z_{k+1}) - \theta_k(z_k)\big) && \text{(since the real parts cancel)} \\
&= \frac{1}{2\pi} \sum_{k=1}^{N} \big(\theta_{k-1}(z_k) - \theta_k(z_k)\big) && \text{(since } z_0 = z_N\text{)} \\
&= n_1 + n_2 + \cdots + n_N,
\end{aligned}$$

which is an integer.

For $k = 0, \ldots, N-1$, define $\eta_k$ on $[t_k, t_{k+1}]$ by $\eta_k := \theta_k \circ \gamma$; as the composite of continuous functions, $\eta_k$ is continuous. We patch together the functions $\eta_k$ to form the function $\eta$ required in (2), adjusting the constants as we go to make $\eta$ continuous at the points $t_k$. The final recipe is

$$\eta(t) = \begin{cases} \eta_0(t) & \text{if } t \in [t_0, t_1], \\ \eta_k(t) + \sum_{r=1}^{k} \big(\eta_{r-1}(t_r) - \eta_r(t_r)\big) & \text{if } t \in [t_k, t_{k+1}] \ \ (1 \leqslant k \leqslant N-1). \end{cases}$$

Finally,

$$\begin{aligned}
2\pi \mathrm{n}(\gamma, 0) &= \sum_{k=0}^{N-1} \big(\theta_k(\gamma(t_{k+1})) - \theta_k(\gamma(t_k))\big) && \text{(from above)} \\
&= \sum_{k=0}^{N-1} \big(\eta_k(t_{k+1}) - \eta_k(t_k)\big) && \text{(by definition of } \eta_k\text{)} \\
&= \sum_{k=0}^{N-1} \big(\eta(t_{k+1}) - \eta(t_k)\big) && \text{(by definition of } \eta\text{)} \\
&= \eta(\beta) - \eta(\alpha). && \square
\end{aligned}$$

**12.10 Continuous selection of argument.** We call the function $\eta$ in Theorem 12.9 a **continuous selection of argument along** $\gamma$. Note that $\eta$ is required to vary continuously with $t$, rather than with $z = \gamma(t)$ (cf. 12.7 and 9.2); when $n(\gamma, 0) \neq 0$, we cannot find a continuous argument function (of $z$) on $\gamma^*$ since no choice from $[\![ \arg z ]\!]$ at $z = \gamma(\alpha) = \gamma(\beta)$ is compatible with continuity; see 9.2.

**12.11 Examples (continuous argument).**

- Let $\gamma(t) = e^{it}$ $(t \in [0, 2\pi])$. For a fixed integer $n$, let $\eta(t) = t + 2n\pi$. This gives a continuous selection of argument along $\gamma$ for any $n$. Once $\eta(0)$ has been decided, the other values of $\eta(t)$ are dictated by the continuity restriction. In particular, we must have $\eta(2\pi) = \eta(0) + 2\pi$.

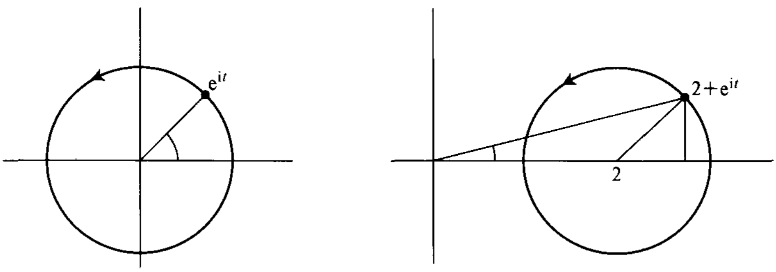



**Figure 12.3** Illustrating argument selection

- Let $\gamma(t) = 2 + e^{it}$ $(t \in [0, 2\pi])$. In this case, a possible choice for $\eta$ is given by
$$\eta(t) = \tan^{-1}\left(\frac{\sin t}{2 + \cos t}\right),$$
where we take the principal value of $\tan^{-1}$ (having values in $(-\pi/2, \pi/2)$). As $t$ increases from 0 to $2\pi$, the value of $\eta(t)$ increases from 0 to $\pi/6$, decreases from $\pi/6$ to $-\pi/6$, and then increases again to its original value 0.

## Cauchy's theorem revisited

Neither Cauchy's theorem I nor Cauchy's theorem II fully reveals what makes the Cauchy theorems work. To clarify matters, we put forward without proof a third, topologically quite sophisticated, Cauchy theorem. An elegant and relatively elementary proof of this result can be found in [**19**].

**12.12 Cauchy's theorem III.** Let $G$ be a region and let $f \in \mathrm{H}(G)$. Then $\int_\gamma f(z)\,\mathrm{d}z = 0$ for any closed path $\gamma$ in $G$ such that $\mathrm{n}(\gamma, w) = 0$ for all $w \notin G$.

What is crucial here is the interaction between $\gamma$ and $G$ via the index. Versions of Cauchy's theorem which do not mention index incorporate geometric restrictions on $\gamma$ or topological restrictions on $G$ which force the index condition to hold. Intuitively, the condition says that $\gamma$ does not wind round points outside $G$. Our comments in 12.8 show that Cauchy's theorem III is a natural generalization of Cauchy's theorem I (for contours). The connection with Cauchy's theorem II is more subtle, and proper appreciation of it demands an understanding of algebraic topology (specifically of the relation between homotopy and homology). Some sense of perspective is conveyed by another deep theorem.

**12.13 Theorem.** Let $G$ be a region. Then the following are equivalent:

(1) $G$ is simply connected;

(2) $\mathrm{n}(\gamma, w) = 0$ for all closed paths in $G$ and for all $w \notin G$;

(3) $\int_\gamma f(z)\,\mathrm{d}z = 0$ for all closed paths $\gamma$ in $G$ and all $f \in \mathrm{H}(G)$;

(4) each $f \in \mathrm{H}(G)$ has an antiderivative (that is, $f = F'$ for some $F \in \mathrm{H}(G)$);

(5) given any $f \in \mathrm{H}(G)$ with $f\colon G \to \mathbb{C} \setminus \{0\}$, there exists a holomorphic logarithm of $f$ (that is, there exists $g \in \mathrm{H}(G)$ such that $\mathrm{e}^g = f$).

The assertion (1) $\implies$ (3) is Cauchy's theorem III. The implication (3) $\implies$ (2) follows from the fact that $f(z) = (z - w)^{-1}$ is holomorphic in $G$ and

$$\int_\gamma f(z)\,\mathrm{d}z = 2\pi \mathrm{i}\,\mathrm{n}(\gamma, w).$$

The implication (3) $\implies$ (4) has already been established, and (4) $\implies$ (5) is an extension of Theorem 12.7. Completing the circle (by proving (5) $\implies$ (1)) is much harder , and well beyond the scope of this book. We refer the interested reader to **[19]**. We also recommend **[10]** to anyone wishing to gain a deeper understanding of index and argument.

## Exercises

12.1 Describe $\gamma^*$ for each of the following closed paths $\gamma$ and use 12.8 to compute $\mathrm{n}(\gamma, 0)$ in each case.

(i) $\gamma$ is the join of $\gamma_1$, $\gamma_2$, and $\gamma_3$, where $\gamma_k(t) := (\frac{5}{2} - k) + ke^{it}$ $(t \in [0, 2\pi])$;

(ii) $\gamma$ is the join of $[-2, -1]$, $-\Gamma_1$, $[1, 2]$, and $\Gamma_2$, where $\Gamma_r(t) = re^{it}$ $(t \in [0, \pi])$;

(iii) $\gamma$ is the join of $[-5, -1]$ and $\gamma_1$, where $\gamma_1(t) = te^{it}/\pi$ $(t \in [\pi, 5\pi])$.

12.2 Let $\gamma_1, \gamma_2 \colon [\alpha, \beta] \to \mathbb{C}$ be closed paths and define $\gamma(t) = \gamma_1(t)\gamma_2(t)$ and $\Gamma(t) = \gamma_1(t) + \gamma_2(t)$ $(y \in [\alpha, \beta])$.

(i) Show that $\gamma$ and $\Gamma$ are closed paths.

(ii) Show that, if $0 \notin \gamma_1^* \cup \gamma_2^*$, then $\mathrm{n}(\gamma, 0) = \mathrm{n}(\gamma_1, 0) + \mathrm{n}(\gamma_2, 0)$.

(iii) Show that if $|\gamma_1(t)| > |\gamma_2(t)|$ for $t \in [\alpha, \beta]$, then $\mathrm{n}(\Gamma, 0) = \mathrm{n}(\gamma_1, 0)$.

12.3 Prove that $\mathbb{C}_\alpha$ is simply connected (see Examples 12.3). Here is a possible strategy. Given a closed path $\gamma$ with $\gamma^* \subseteq \mathbb{C}_\alpha$,

(i) find a point $a$ such that $[b, z] \subseteq \mathbb{C}_\alpha$ for all $z \in \gamma^*$ and for all $b \in \mathrm{D}(a; r)$ (for some $r > 0$);

(ii) show, with the aid of 4.5, that there exist finitely many open convex sets $G_0, \ldots, G_{N-1}$ such that $a \in \bigcap_{k=0}^{N-1} G_k$ and $\gamma$ is the join of paths $\gamma_k$ for $k = 0, \ldots, N-1$ and $\gamma_k^* \subseteq G_k$ (hint: each $G_k$ may be taken to be an open wedge bounded by an arc and two line segments meeting at a point in $\mathrm{D}(a; r)$);

(iii) deduce that $\gamma$ is homotopic to the null path with image $\{a\}$.

# 13 Cauchy's formulae

Armed with Cauchy's theorem we can prove a host of striking results about holomorphic functions. These stem from the Cauchy formulae which we derive in this chapter. We are then able to prove the following, with relative ease.

- **Liouville's theorem** A function which is holomorphic in $\mathbb{C}$ cannot be bounded unless it is constant (13.3).
- **Infinite differentiability** Any holomorphic function (assumed to be differentiable just once) is in fact automatically infinitely differentiable (Theorem 13.7).
- **Taylor's theorem** Any holomorphic function is locally representable by power series (14.4).
- **Identity theorem (Corollary)** If $f$ is holomorphic in a region $G$ and is zero in an open disc within $G$ then $f$ is identically zero in $G$ (15.8).

All this is in sharp contrast to the behaviour of real-valued functions on $\mathbb{R}$. The contrast is very welcome: the theorems are not hedged around with unmemorable technical restrictions.

## Cauchy's integral formula

Cauchy's integral formula expresses the value of a holomorphic function at a point $a$ in terms of a 'boundary value integral' taken round a contour encircling the point. The ingredients in the proof are

- deformation, which allows us to replace the given contour by a small circle;
- the fact that $\int_{\gamma(a;r)}(w-a)^{-1}\,\mathrm{d}w = 2\pi\mathrm{i}$;
- estimation, exploiting the fact that $f(w) - f(a) \to 0$ as $w \to a$.

We have elected to use $w$ rather than $z$ as our dummy variable of integration because we shall in due course want to rename $a$ as $z$ when we want to treat this as a variable.

In Cauchy's theorem, the orientation of the contour did not need to be specified. In Cauchy's integral formula, and all subsequent results giving formulae for integrals whose values are not in general zero, the contour is taken to be positively oriented.

**13.1 Cauchy's integral formula.** Let $f$ be holomorphic inside and on a positively oriented contour $\gamma$. Then, if $a$ is inside $\gamma$,

$$f(a) = \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{w-a}\, \mathrm{d}w.$$

**Proof** There exists $R$ such that $\mathrm{D}(a;R) \subseteq \mathbf{I}(\gamma)$. For any $r < R$,

$$\int_\gamma \frac{f(w)}{w-a}\, \mathrm{d}w = \int_{\gamma(a;r)} \frac{f(w)}{w-a}\, \mathrm{d}w,$$

by the Deformation theorem (11.9(1)). Also, because $f(a)$ is constant,

$$\int_{\gamma(a;r)} \frac{f(a)}{w-a}\, \mathrm{d}w = f(a)\int_{\gamma(a;r)} \frac{1}{w-a}\, \mathrm{d}w = 2\pi \mathrm{i} f(a)$$

from the Fundamental integral (10.4). Now we calculate the difference between the integral in the formula and the desired value, and estimate:

$$\begin{aligned} \left| \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{w-a}\, \mathrm{d}w - f(a) \right| &= \left| \frac{1}{2\pi \mathrm{i}} \int_{\gamma(a;r)} \frac{f(w)-f(a)}{w-a}\, \mathrm{d}w \right| \\ &= \left| \frac{1}{2\pi \mathrm{i}} \int_0^{2\pi} \frac{f(a+r\mathrm{e}^{\mathrm{i}\theta}) - f(a)}{r\mathrm{e}^{\mathrm{i}\theta}}\, \mathrm{i}r\mathrm{e}^{\mathrm{i}\theta}\, \mathrm{d}\theta \right| \\ &\leqslant \frac{1}{2\pi} 2\pi \times \sup_{\theta \in [0,2\pi]} \left| f(a+r\mathrm{e}^{\mathrm{i}\theta}) - f(a) \right|, \end{aligned}$$

by the Estimation theorem (10.10). Since $f$ is continuous at $a$ (recall 5.11), the supremum above tends to zero as $r \to 0$. The expression on the left-hand side of the display is independent of $r$ and so must be zero. □

**13.2 Examples (Cauchy's integral formula).**

- $\displaystyle\int_{\gamma(4;5)} \frac{\cos z}{z}\, \mathrm{d}z = 2\pi \mathrm{i}\big[\cos z\big]_{z=0} = 2\pi \mathrm{i}.$
- $\displaystyle\int_{\gamma(\mathrm{i};1)} \frac{z^2}{z^2+1}\, \mathrm{d}z = 2\pi \mathrm{i}\big[z^2/(z+\mathrm{i})\big]_{z=\mathrm{i}} = -\pi.$
- We cannot evaluate

$$I := \int_{\gamma(0;2)} \frac{\mathrm{e}^{\mathrm{i}\pi z/2}}{(z^2-1)}\, \mathrm{d}z$$

directly by Cauchy's integral formula because the integrand fails to be holomorphic at two points inside $\gamma(0;2)$, namely 1 and $-1$. However a partial fraction decomposition allows us to write

$$\begin{aligned} I &= \frac{1}{2}\int_{\gamma(0;2)} \frac{e^{i\pi z/2}}{z-1}\,dz - \frac{1}{2}\int_{\gamma(0;2)} \frac{e^{i\pi z/2}}{z+1}\,dz \\ &= \left[\frac{1}{2}e^{i\pi z/2}\right]_{z=1} - \left[\frac{1}{2}e^{i\pi z/2}\right]_{z=-1} \\ &= i. \end{aligned}$$

Use of partial fractions is a feasible method in the last example but it is laborious. Later we shall have more powerful and more effective methods for evaluating integrals like this one, and many other integrals too.

We observed in 7.9 that $\cos z$ is not bounded in $\mathbb{C}$. Liouville's theorem shows that this behaviour is typical of non-constant functions which are holomorphic everywhere.

**13.3 Liouville's theorem (via Cauchy's integral formula).** Let $f$ be holomorphic and bounded in the complex plane $\mathbb{C}$. Then $f$ is constant.

**Proof** Suppose that $|f(w)| \leqslant M$ for all $w \in \mathbb{C}$. Fix $a$ and $b$ in $\mathbb{C}$. Take $R \geqslant 2\max\{|a|, |b|\}$, so that $|w-a| \geqslant \frac{1}{2}R$ and $|w-b| \geqslant \frac{1}{2}R$ whenever $|w| = R$ (by 1.9(3)). By Cauchy's integral formula applied with $\gamma = \gamma(0;R)$,

$$\begin{aligned} f(a) - f(b) &= \frac{1}{2\pi i}\int_{\gamma} f(w)\left(\frac{1}{w-a} - \frac{1}{w-b}\right) dw \\ &= \frac{a-b}{2\pi i}\int_{\gamma} \frac{f(w)}{(w-a)(w-b)}\,dw, \end{aligned}$$

so, by the Estimation theorem,

$$|f(a) - f(b)| \leqslant \frac{1}{2\pi} 2\pi R M \frac{|a-b|}{(\frac{1}{2}R)^2}.$$

The right-hand side can be made arbitrarily small by taking $R$ sufficiently large. Hence $f(a) = f(b)$ for any $a$ and $b$ in $\mathbb{C}$. □

Liouville's theorem yields an unexpected bonus: an easy proof of the famous result commonly known as the Fundamental theorem of algebra.

**13.4 Fundamental theorem of algebra.** Let $p(z)$ be a non-constant polynomial with complex coefficients. Then there exists $\zeta \in \mathbb{C}$ such that $p(\zeta) = 0$. Consequently, a complex polynomial of degree $n > 1$ has $n$ roots (not necessarily distinct) in $\mathbb{C}$.

**Proof** Suppose for a contradiction that $p(z) \neq 0$ for every $z$. Since $|p(z)| \to \infty$ as $|z| \to \infty$, there exists $R$ such that $|1/p(z)| < 1$ for $|z| > R$. On the compact set $\overline{\mathrm{D}}(0; R)$, $1/p(z)$ is continuous and hence bounded, by 3.24. Hence $1/p(z)$ is bounded on $\mathbb{C}$. It is also holomorphic (see 5.8), and so must be constant, by Liouville's theorem. We have the required contradiction. The final statement is proved by induction on the degree of the polynomial. □

## Higher-order derivatives

**13.5 Onwards from Cauchy's integral formula.** Let $f$ be holomorphic inside and on a positively oriented contour $\gamma$. Then

(CIF-0) $$f(z) = \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{(w-z)}\,\mathrm{d}w \qquad (z \in \mathbf{I}(\gamma)).$$

From this we would hope to be able to deduce

(CIF-1) $$f'(z) = \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{(w-z)^2}\,\mathrm{d}w \qquad (z \in \mathbf{I}(\gamma)),$$

by differentiating the right-hand side of (CIF-0) with respect to $z$. We say 'hope' because this presumes that differentiation and integration can legitimately be interchanged. In 13.6 we show that this is indeed so. (Each of integration and differentiation is defined in terms of a limiting process; it is a well-known fact in analysis that taking repeated limits in different orders may give different values.)

We may try to take this further by differentiating $\int_\gamma f(w)/(w-z)^2\,\mathrm{d}w$ with respect to $w$, in the hope of obtaining

(CIF-2) $$f''(z) = \frac{2}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{(w-z)^3}\,\mathrm{d}w \qquad (z \in \mathbf{I}(\gamma)).$$

There is a significant difference between this and the first differentiation. Since $f$ is assumed to be holomorphic we know that $f'(z)$ exists. We do *not* know in advance that $f''(z)$ exists. But we can use (CIF-1) to investigate the quotient $(f'(z+h) - f'(z))/h$ with a view to showing that its limit as $h \to 0$ exists and is given by the right-hand side of (CIF-2). This can be validated. The process can then be repeated to show, successively, the existence of $f^{(n)}(z)$ for $n = 3, 4, \ldots$, all given by differentiation under the integral sign:

(CIF-n) $$f^{(n)}(z) = \frac{n!}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{(w-z)^{n+1}}\,\mathrm{d}w \qquad (z \in \mathbf{I}(\gamma)).$$

We call the right-hand side of (CIF-n) a **Cauchy integral.**

We give below a detailed justification of (CIF-1), to validate differentiation of an integral in a simple case. This is instructive as an illustration of techniques for manipulating and estimating integrals of this type. We then outline a proof of (CIF-2). This result leads to the important result that holomorphic functions are infinitely differentiable. We also present an inductive proof for (CIF-n) for general $n$ (13.9). This somewhat forbidding proof, which overrides those for both (CIF-1) and (CIF-2), may be skipped by anyone content to take it on trust. [A proof based on the Lebesgue Dominated convergence theorem is slicker.]

**13.6 Cauchy's formula for the first derivative.** Let $f$ be holomorphic inside and on a positively oriented contour $\gamma$ and let $a$ be inside $\gamma$. Then

$$f'(a) = \frac{1}{2\pi \mathrm{i}} \int_{\gamma} \frac{f(w)}{(w-a)^2}\,\mathrm{d}w.$$

**Proof** As in the proof of Cauchy's integral formula, we may use the Deformation theorem to replace the integral by the corresponding integral round a circle, $\gamma(a; 2r)$ say. Cauchy's integral formula gives

$$\begin{aligned}\frac{f(a+h)-f(a)}{h} &= \frac{1}{2h\pi \mathrm{i}} \int_{\gamma(a;2r)} f(w) \left(\frac{1}{w-a-h} - \frac{1}{w-a}\right) \mathrm{d}w \\ &= \frac{1}{2\pi \mathrm{i}} \int_{\gamma(a;2r)} \frac{f(w)}{(w-a-h)(w-a)}\,\mathrm{d}w.\end{aligned}$$

Hence

$$\begin{aligned}\frac{f(a+h)-f(a)}{h} - \frac{1}{2\pi \mathrm{i}} \int_{\gamma(a;2r)} \frac{f(w)}{(w-a)^2}\,\mathrm{d}w &\\ = \frac{1}{2\pi \mathrm{i}} \int_{\gamma(a;2r)} f(w) \left(\frac{1}{(w-a-h)(w-a)} - \frac{1}{(w-a)^2}\right) \mathrm{d}w &\\ = \frac{h}{2\pi \mathrm{i}} \int_{\gamma(a;2r)} \frac{f(w)}{(w-a-h)(w-a)^2}\,\mathrm{d}w. &\end{aligned}$$

We claim that the last expression tends to 0 as $h \to 0$. This looks plausible: we just need to find a constant which acts as an upper bound for the modulus of the integral when $|h|$ is small. It seems likely that this integral is bounded so long as $|w-a-h|$ is bounded away from zero.

Here is the detailed proof of the claim. Choose $h$ such that $|h| < r$ so that, by 1.9(3), $|w-a-h| \geqslant |w-a| - |h| > r$ for all $w \in \gamma(a; 2r)^*$. Also, since $f$ is continuous on $\gamma(a; 2r)^*$, which is compact, there exists a constant $M$ such that

$|f(w)| \leqslant M$ for all $w$ with $|w-a| = 2r$ (3.24). By the Estimation theorem,

$$\left| \frac{f(a+h)-f(a)}{h} - \frac{1}{2\pi \mathrm{i}} \int_{\gamma(a;2r)} \frac{f(w)}{(w-a)^2}\, \mathrm{d}w \right| \leqslant \frac{|h|}{2\pi} \frac{M}{4r^3}$$
$$\to 0 \text{ as } h \to 0. \qquad \square$$

Note that the result below can be derived without use of the Deformation theorem, since its proof uses 13.1 and 13.6 only in the case that the contour $\gamma$ is a circle centred on a point $a$. The result asserts in particular that the derivative of a holomorphic function is holomorphic.

**13.7 Theorem (the existence of derivatives).** Suppose that $f$ is holomorphic in an open set $G$. Then

(1) $f' \in H(G)$;

(2) $f$ has derivatives of all orders in $G$.

**Proof** Fix $a \in G$ and choose $r > 0$ such that $\overline{\mathrm{D}}(a;2r) \subseteq G$. For $|h| < r$, (CIF-1) gives

$$\frac{f'(a+h)-f'(a)}{h} = \frac{1}{2\pi \mathrm{i}} \int_{\gamma(a;2r)} f(w) \left( \frac{1}{(w-a-h)^2} - \frac{1}{(w-a)^2} \right) \mathrm{d}w.$$

The right-hand side can be shown to tend to $2\int_{\gamma(a;2r)} (f(w)/(w-a)^3)\, \mathrm{d}w$ by an argument similar to that used in 13.6.

Thus $f''(a)$ exists for all $a \in G$, so that $f' \in \mathrm{H}(G)$. Replacing $f$ by $f'$, we see that this in turn implies that $f'''(a)$ exists for all $a \in G$. An inductive argument now shows that $f^{(n)}$ exists throughout $G$, for all $n$. $\square$

One consequence of 13.7(1) is a partial converse to Cauchy's theorem.

**13.8 Morera's theorem.** Suppose that $f$ is continuous on an open set $G$ and that $\int_\gamma f(z)\, \mathrm{d}z = 0$ for all triangles $\gamma$ in $G$. Then $f \in \mathrm{H}(G)$.

**Proof** Let $a \in G$ and choose $r$ such that $\mathrm{D}(a;r) \subseteq G$. Since $\mathrm{D}(a;r)$ is a convex region, the Indefinite integral theorem I (11.3) supplies $F \in \mathrm{H}(\mathrm{D}(a;r))$ such that $F' = f$. By Theorem 13.7, $f \in \mathrm{H}(\mathrm{D}(a;r))$. Since $a$ is arbitrary, $f \in \mathrm{H}(G)$. $\square$

We have proved that a holomorphic function has derivatives of all orders and have obtained an integral formula for the first derivative. We now provide the promised justification of the corresponding formula for any derivative.

**13.9 Cauchy's formula for derivatives.** Let $f$ be holomorphic inside and on a positively oriented contour $\gamma$ and let $a$ lie inside $\gamma$. Then $f^{(n)}(a)$ exists for $n = 1, 2, \ldots$ and

$$\text{(CIF-n)} \qquad f^{(n)}(a) = \frac{n!}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{(w-a)^{n+1}}\, \mathrm{d}w \qquad (a \in \mathbf{I}(\gamma)).$$

**Proof** We proceed by induction. The base case can be taken to be $n = 0$, for which the result is Cauchy's integral formula. Assume (CIF-k) holds for all $a \in \mathbf{I}(\gamma)$. We shall prove (CIF-(k+1)). By the Deformation theorem, we may assume that $\gamma = \gamma(a; 2r)$ for some $r > 0$. The argument below differs from that used already in the special cases $k = 0, 1$ only in its backwards use of the Fundamental theorem of calculus to handle differences. Take $|h| < r$. By (CIF-k),

$$\begin{aligned} f^{(k+1)}(a+h) - f^{(k)}(a) &= \frac{k!}{2\pi \mathrm{i}} \int_\gamma f(w) \left( \frac{1}{(w-a-h)^{k+1}} - \frac{1}{(w-a)^{k+1}} \right) \mathrm{d}w \\ &= \frac{(k+1)!}{2\pi \mathrm{i}} \int_\gamma f(w) \left( \int_{[a,a+h]} (w-\zeta)^{-k-2}\, \mathrm{d}\zeta \right) \mathrm{d}w, \end{aligned}$$

by the Fundamental theorem of calculus.

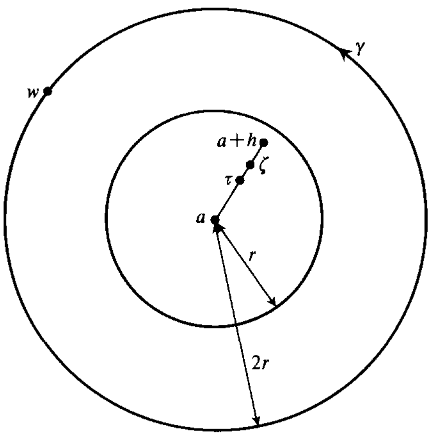



**Figure 13.1** Proof of (CIF-n)

We shall show that $F(h) \to 0$ as $h \to 0$, where

$$F(h) = \frac{f^{(k+1)}(a+h) - f^{(k)}(a)}{h} - \frac{(k+1)!}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{(w-a)^{k+2}}\, \mathrm{d}w$$

$$= \frac{(k+1)!}{2\pi i h}\int_\gamma f(w)\Big(\int_{[a,a+h]}\big((w-\zeta)^{-k-2}-(w-a)^{-k-2}\big)\,d\zeta\Big)\,dw$$
$$= \frac{(k+2)!}{2\pi i h}\int_\gamma f(w)\Big[\int_{[a,a+h]}\Big(\int_{[a,\zeta]}(w-\tau)^{-k-3}\,d\tau\Big)\,d\zeta\Big]\,dw.$$

Since $f$ is holomorphic, and so continuous, it is bounded, by $M$ say, on the compact set $\gamma^*$. For $\tau \in [a,\zeta]$ and $\zeta \in [a, a+h]$, we have $|w-\tau| \geqslant r$ for all $w \in \gamma^*$. Also, $|\zeta - a| \leqslant r$ whenever $|w-a| = 2r$. Also $|\zeta - a| \leqslant |h|$. See Fig. 13.1. By the Estimation theorem,

$$|F(h)| \leqslant \frac{(k+2)!}{2\pi\,|h|} \times \frac{M\,|h|^2}{r^{k+3}} \times 4\pi r,$$

so $F(h) \to 0$ as $h \to 0$. □

**13.10 Example (Cauchy's formulae for derivatives).** By 13.9,

- $\displaystyle\int_{\gamma(0;1)} e^z z^{-3}\,dz = \left[\frac{2\pi i}{2!}\frac{d^2}{dz^2}e^z\right]_{z=0} = \pi i,$
- $\displaystyle\int_{\gamma(1;5/2)} \frac{1}{(z-4)(z+1)^4}\,dz = \left[\frac{2\pi i}{3!}\frac{d^3}{dz^3}(z-4)^{-1}\right]_{z=-1} = -\frac{2\pi i}{5^4}.$

**13.11 Stocktaking.** We conclude this chapter with a summary of how our repertoire of techniques for evaluating integrals has been enlarged by the Cauchy formulae. We now know that, given a positively oriented contour $\gamma$, a point $a \notin \gamma^*$, and a function $f$ holomorphic inside and on $\gamma$, we have, for $n = 0, 1, \ldots,$

$$\int_\gamma \frac{f(z)}{(z-a)^{n+1}}\,dz = \begin{cases} 2\pi i f(a) & \text{if } n = 0 \text{ and } a \in \mathbf{I}(\gamma) \quad \text{(by 13.3)},\\ \dfrac{2\pi i}{n!}f^{(n)}(a) & \text{if } n \geqslant 1 \text{ and } a \in \mathbf{I}(\gamma) \quad \text{(by 13.9)},\\ 0 & \text{if } a \in \mathbf{O}(\gamma) \quad \text{(by 11.6)}.\end{cases}$$

(The case in which $f$ is the constant function 1 was covered in 11.10.)

## Exercises

13.1 Evaluate, when $\gamma = \gamma(0;2)$,

$$\text{(i)} \int_\gamma \frac{z^3+5}{z-\mathrm{i}}\,\mathrm{d}z, \quad \text{(ii)} \int_\gamma \frac{1}{z^2+z+1}\,\mathrm{d}z, \quad \text{(iii)} \int_\gamma \frac{\sin z}{z^2+1}\,\mathrm{d}z.$$

13.2 Evaluate $\int_{\gamma(0;1)} z^{-1}\cos z\,\mathrm{d}z$. By writing the integral in parametric form, deduce that

$$\int_0^{2\pi} \cos(\cos\theta)\cosh(\sin\theta)\,\mathrm{d}\theta = 2\pi.$$

13.3 Let $f$ be holomorphic inside and on $\gamma(0;1)$. Prove that

$$2\pi\mathrm{i}f(z) = \int_{\gamma(0;1)} \frac{f(w)}{w-z}\,\mathrm{d}w - \int_{\gamma(0;1)} \frac{f(w)}{w-1/\overline{z}}\,\mathrm{d}w \quad (0<|z|<1).$$

Hence prove the **Poisson integral formula**

$$f(re^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{(1-r^2)}{(1-2r\cos(\theta-t)+r^2)} f(e^{\mathrm{i}t})\,\mathrm{d}t \quad (0<r<1).$$

13.4 Explain why

$$\int_{\gamma(0;1)} \frac{\operatorname{Re} z}{(z-\frac{1}{2})}\,\mathrm{d}z$$

cannot be evaluated by applying Cauchy's integral formula with $f(z) = \operatorname{Re} z$. Prove that $\operatorname{Re} z$ coincides with $(z+z^{-1})/2$ when $|z|=1$. Hence evaluate the given integral.

13.5 By considering the complex conjugate of its parametric form, evaluate the integral

$$\frac{1}{2\pi\mathrm{i}}\int_{\gamma(0;1)} \frac{\overline{f(z)}}{z-a}\,\mathrm{d}z$$

in the cases (i) $|a|<1$, (ii) $|a|>1$, where $f \in \mathrm{H}(\mathrm{D}(0;R))$ $(R>1)$.

13.6 Suppose that $f$ is holomorphic in $\mathbb{C}$ and such that $f(z) = f(z+2\pi)$ and $f(z) = f(z+2\pi\mathrm{i})$ for all $z \in \mathbb{C}$. Use Liouville's theorem to prove that $f$ is constant. (Hint: consider the restriction of $f$ to suitable squares.)

13.7 Let $f$ be holomorphic in $\mathbb{C}$ and such that $f(z) \to 0$ as $|z| \to \infty$. Prove that $f$ is identically zero.

13.8 Let $f$ be holomorphic in $\mathbb{C}$.

(i) Prove that if $|f(z)| > M$ in $\mathbb{C}$ then $f$ is constant.

(ii) Prove that if $e^f$ is bounded then $f$ is constant.

(iii) Prove that if $\operatorname{Re} f$ is bounded either above or below then $f$ is constant. (Hint: consider suitable exponentials.)

(You will need the result of 7.4.)

13.9 Evaluate, when $\gamma = \gamma(0;2)$,

$$\text{(i)} \int_\gamma (z-1)^{-3} e^{z^2}\, dz, \qquad \text{(ii)} \int_\gamma z^{-n} \cos z\, dz \quad (n = 1, 2, \ldots),$$

$$\text{(iii)} \int_\gamma \frac{1}{(z+1)^2(z^2+9)}\, dz.$$

13.10 Evaluate $\displaystyle\int_{\gamma(0;1)} z^n (1-z)^m\, dz \quad (m = 0, 1, 2, \ldots,\ n = 0, \pm 1, \pm 2, \ldots)$.

13.11 Let $a, b \in \mathbb{C}$ with $|a| \neq 1$, $|b| \neq 1$. Evaluate, distinguishing cases,

$$\text{(i)} \int_{\gamma(0;1)} \frac{1}{(z-a)(z-b)}\, dz, \qquad \text{(ii)} \int_{\gamma(0;1)} \left(\frac{z-b}{z-a}\right)^2 dz.$$

13.12 Suppose $G$ is an open set, $f\colon G \to \mathbb{C}$ is continuous, and $f$ is holomorphic in $G \setminus [a, b]$. Use Morera'a theorem to prove that $f \in H(G)$.

# 14 Power series representation

In Chapter 7 we proved the major result that a convergent power series defines a holomorphic function. This chapter reveals in full the symbiotic relationship that exists between holomorphic functions and power series. We begin with some technical results, needed for manipulating integrals and series.

## Integration of series in general and power series in particular

**14.1 Interchange of summation and integration.** Suppose that $\gamma$ is a path and that $u_0, u_1, \ldots$ are continuous (or piecewise continuous) functions on $\gamma^*$. It is certainly true that, for any natural number $N$,

$$\sum_{k=0}^{N} \int_{\gamma} u_k(z)\,\mathrm{d}z = \int_{\gamma} \sum_{k=0}^{N} u_k(z)\,\mathrm{d}z.$$

If the finite sum here is replaced by an infinite sum, the corresponding interchange of summation and integration may well not be valid. For integrals along paths, a systematic study of sufficient conditions for interchange to be valid can be based either on a basic treatment of uniform convergence or on more sophisticated techniques. Since we need only to be able to handle the integration of series closely related to power series, taken round quite special contours, we elect on the basic track to avoid uniform convergence and to adopt an *ad hoc* approach. An optional 'minimum kit' treatment of uniform convergence is given at the end of the chapter.

The use of Theorem 14.2 is illustrated below in 14.3 and in the proof of Theorem 14.4.

**14.2 Interchange theorem (simple form).** Let $\gamma$ be a path, let $U, u_0, u_1, \ldots$ be continuous on $\gamma^*$, and assume that $\sum_{k=0}^{\infty} u_k(z)$ converges to $U(z)$ for all

$z \in \gamma^*$. Assume that there exist constants $M_k$ such that $\sum M_k$ converges and $|u_k(z)| \leqslant M_k$ for all $z \in \gamma^*$. Then

$$\sum_{k=0}^{\infty} \int_\gamma u_k(z)\,\mathrm{d}z = \int_\gamma \sum_{k=0}^{\infty} u_k(z)\,\mathrm{d}z = \int_\gamma U(z)\,\mathrm{d}z.$$

[If you already know about uniform convergence you will recognize Weierstrass' $M$-test lurking here. You should also realize that our hypotheses are stronger than necessary: continuity of $U$ is automatic from the other conditions.]

**Proof** For $N = 0, 1, \ldots$, let $U_N(z) = \sum_{k=0}^{N} u_k(z)$. Both $U_N$ and $U$ are continuous, and hence integrable, on $\gamma^*$. Also, by 6.1(3), $\sum |u_k(z)|$ converges. We now have, for fixed $N$,

$$\begin{aligned}
&\left| \int_\gamma U(z)\,\mathrm{d}z - \sum_{k=0}^{N} \int_\gamma u_k(z)\,\mathrm{d}z \right| \\
&\qquad = \left| \int_\gamma \big(U(z) - U_N(z)\big)\,\mathrm{d}z \right| \\
&\qquad \leqslant \sup_{z\in\gamma^*} \{|U(z) - U_N(z)|\} \times \operatorname{length}(\gamma) \qquad \text{(by 10.10)} \\
&\qquad \leqslant \sup_{z\in\gamma^*} \sum_{k=N+1}^{\infty} |u_k(z)| \times \operatorname{length}(\gamma) \qquad \text{(see Exercise 6.1)} \\
&\qquad \leqslant \sum_{k=N+1}^{\infty} M_k \times \operatorname{length}(\gamma)
\end{aligned}$$

and this tends to zero as $N \to \infty$, because $\sum M_k$ converges. □

**14.3 The coefficients in a power series.** Let $f(z) = \sum_{k=0}^{\infty} c_k z^k$, where the power series has radius of convergence $R > 0$. We claim that

$$c_n = \frac{1}{2\pi \mathrm{i}} \int_{\gamma(0;r)} \frac{f(z)}{z^{n+1}}\,\mathrm{d}z \quad (0 \leqslant r < R,\ n = 0, 1, \ldots).$$

Provided summation and integration can be interchanged we have, for fixed $n$ and fixed $r < R$,

$$\begin{aligned}
\int_{\gamma(0;r)} \frac{f(z)}{z^{n+1}}\,\mathrm{d}z &= \int_{\gamma(0;r)} \left( \sum_{k=0}^{\infty} c_k z^k \right) z^{-n-1}\,\mathrm{d}z \\
&= \sum_{k=0}^{\infty} c_k \int_{\gamma(0;r)} z^{k-n-1}\,\mathrm{d}z \\
&= 2\pi \mathrm{i} c_n \qquad \text{(by 10.4).}
\end{aligned}$$

We justify this by applying Theorem 14.2 with $\gamma = \gamma(0;r)$, $u_k(z) = c_k z^{k-n-1}$, and $U(z) = z^{-n-1}f(z)$; by 6.8(2), $U$ is continuous. On $\gamma(0;r)^*$, we have $|u_k(z)| = M_k := |c_k|\, r^{k-n-1}$, and $\sum M_k$ converges by the Radius of convergence lemma (6.6). □

## Taylor's theorem

We now use Cauchy's integral formula to prove that any function holomorphic in a disc has a power series expansion. It is thanks to Cauchy's theorem that Theorem 14.4 is stronger and more satisfactory than most forms of Taylor's theorem for a function of a real variable.

**14.4 Taylor's theorem.** Let $f \in \mathrm{H}(\mathrm{D}(a;R))$ $(R > 0)$. Then there exist unique constants $c_n$ such that

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n \quad (z \in \mathrm{D}(a;R)).$$

The constant $c_n$ is given by

$$c_n = \frac{1}{2\pi\mathrm{i}} \int_\gamma \frac{f(w)}{(w-a)^{n+1}}\,\mathrm{d}w = \frac{f^{(n)}(a)}{n!},$$

where $\gamma$ is a circle $\gamma(a;r)$ $(0 < r < R)$ or is any positively oriented contour $\tilde{\gamma}$ in $\mathrm{D}(a;R)$ enclosing $a$, where $\overline{\mathrm{D}}(a;r) \subseteq \mathbf{I}(\gamma)$ [or any simple closed path homotopic in $\mathrm{D}'(a;R)$ to $\gamma$].

**Proof** Fix $z \in \mathrm{D}(a;R)$ and choose $r$ such that $|z-a| < r < R$. Take $\gamma = \gamma(a;r)$. By Cauchy's integral formula,

$$f(z) = \frac{1}{2\pi\mathrm{i}} \int_\gamma \frac{f(w)}{w-z}\,\mathrm{d}w.$$

Since $|z-a| < |w-a|$ for all $w \in \gamma^*$, the right-hand side of the equation

$$\frac{1}{w-z} = \frac{1}{w-a}\,\frac{1}{(1-((z-a)/(w-a))}$$

can be expanded binomially (see 6.3) to give

$$f(z) = \frac{1}{2\pi\mathrm{i}} \int_\gamma \sum_{n=0}^{\infty} \frac{(z-a)^n}{(w-a)^{n+1}} f(w)\,\mathrm{d}w.$$

On the compact set $\gamma^*$, the continuous function $f$ is bounded, so for some constant $M$ we have

$$\left|\frac{(z-a)^n}{(w-a)^{n+1}}f(w)\right| \leqslant M_n := \frac{M}{r}\left(\frac{|z-a|}{r}\right)^n.$$

The series $\sum M_n$ converges, since $|z-a| < r$. Hence, by Theorem 14.2 [uniform convergence], summation and integration may be interchanged, to give

$$f(z) = \sum_{n=0}^{\infty}\left(\frac{1}{2\pi \mathrm{i}}\int_\gamma \frac{f(w)}{(w-a)^{n+1}}\,\mathrm{d}w\right)(z-a)^n.$$

The remaining assertions of the theorem now follow from Cauchy's formula for derivatives 13.9 and, for uniqueness, 6.8(4) or 14.3. □

**14.5 Examples (finding Taylor expansions).**

- ⊙ **Tactical tip** We have two possible strategies for computing the Taylor series of a function $f$ about a point $a$. The first is to compute the derivatives $f^{(n)}(a)$. It is rarely to be recommended. The second relies on making use of known expansions and appealing to uniqueness.
- In 6.2 and 6.9 we derived various expansions from the geometric series. All these expansions are the Taylor series of the functions they represent.
- Let $f(z) = z^5 \sin 2z$ in $\mathbb{C}$. It would be arduous to calculate derivatives. However, using the sine series, we obtain

  $$f(z) = \sum_{n=0}^{\infty}(-1)^n 2^{2n+1}\frac{z^{2(n+3)}}{(2n+1)!} \quad (z \in \mathbb{C}).$$

  By uniqueness, this is the Taylor series.

**14.6 Taylor expansion for a holomorphic branch of the logarithm.** Cut the plane along $(-\infty, 0]$ and let $f$ be the holomorphic branch of the logarithm in $\mathbb{C}_\pi = \mathbb{C} \setminus (-\infty, 0]$ given by

$$f(z) = \log|z| + \mathrm{i}\theta \quad (0 \neq z = |z|\,\mathrm{e}^{\mathrm{i}\theta},\ -\pi < \theta < \pi).$$

We know that $f \in \mathrm{H}(\mathrm{D}(1;1))$ and so $f$ must have a Taylor expansion

$$\sum_{n=0}^{\infty} c_n(z-1)^n$$

in the disc $\mathrm{D}(1;1)$. We now find this expansion. Note that

$$\frac{1}{z} = \frac{1}{1+(z-1)} = \sum_{n=0}^{\infty}(-1)^n(z-1)^n \quad (|z-1|<1).$$

By the Differentiation theorem for power series (6.8),

$$\frac{1}{z} = \frac{\mathrm{d}}{\mathrm{d}z}\log z = \sum_{n=1}^{\infty} nc_n(z-1)^{n-1} \quad (|z-1|<1).$$

By uniqueness of power series expansions, $c_n = (-1)^{n-1}1/n$ for $n \geqslant 1$.

The value of $c_0$ is fixed by the choice of branch: We have $f(1) = 0$, so $c_0 = 0$. Therefore

$$f(z) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(z-1)^n \quad (z \in \mathrm{D}(1;1)).$$

Similar arguments yield Taylor expansions for the function $f$ in any disc in $\mathbb{C}_\pi$.

**14.7 Estimating Cauchy integrals.** It will often be important to have an estimate of the magnitude of the integrals

$$c_n = \frac{1}{2\pi \mathrm{i}}\int_{\gamma(0;r)} \frac{f(w)}{w^{n+1}}\,\mathrm{d}w$$

giving the coefficients of the Taylor series of $f \in \mathrm{H}(\mathrm{D}(0;R))$ $(n = 0,1,\ldots, r < R)$. By the Estimation theorem,

$$\begin{aligned}
|c_n| &= \left|\frac{1}{2\pi \mathrm{i}}\int_{\gamma(0;r)} \frac{f(w)}{w^{n+1}}\,\mathrm{d}w\right| \\
&\leqslant \frac{1}{2\pi}\sup\{\,\left|f(z)z^{-n-1}\right| : |z| = r\,\} \times \mathrm{length}(\gamma(0;r)) \\
&= \frac{1}{2\pi}M(r)r^{-n-1} \times 2\pi r \\
&= r^{-n}M(r),
\end{aligned}$$

where $M(r) := \sup\{\,|f(z)| : |z| = r\,\}$.

We can immediately derive an interesting and perhaps surprising consequence of the estimate above.

**14.8 Theorem (forcing a holomorphic function to be a polynomial).** Let $f$ be holomorphic in $\mathbb{C}$, with Taylor expansion $f(z) = \sum_{n=0}^{\infty} c_n z^n$ valid for all $z \in \mathbb{C}$. Suppose that there exist positive constants $M$ and $K$ and a positive integer $k$ such that

$$|f(z)| \leqslant M\,|z|^k \quad (|z| \geqslant K).$$

Then $f$ is a polynomial of degree at most $k$.

**Proof** Let $r \geqslant K$ and note that $M(r) \leqslant Mr^k$. Estimating as in 14.7 we obtain

$$|c_n| \leqslant \frac{1}{2\pi} M r^{k-n-1} \times 2\pi r.$$

Since $r$ can be chosen arbitrarily large, we must have $c_n = 0$ for all $n > k$. Hence $f$ is a polynomial of degree not greater than $k$. □

⊙ Students often regard it as 'obvious' that the coefficients $c_n$ in a Taylor series $\sum c_n z^n$ must be zero for $n > k$ if $|f(z)|$ grows no faster than $|z|^k$. We dispute that it is obvious. But the point is not worth arguing over since the estimation of the integral for the *individual* coefficient $c_n$ gives precise information about its magnitude in terms of information about $f$—without any hand-waving.

**14.9 The role of power series: summing up.** We have now completed our presentation of an important circle of ideas. In Chapter 6 we showed that every power series $\sum_{n=0}^{\infty} c_n (z-a)^n$ with radius of convergence $R > 0$ defines a holomorphic function in $\mathrm{D}(a; R)$. In the opposite direction, Taylor's theorem shows that every function holomorphic in an open set $G$ is **analytic**, that is, locally representable by power series.

It is possible to develop much of complex function theory directly in terms of analytic functions. Starting from a power series, 14.3 gives us the coefficients as Cauchy integrals. Combining 14.3 with Theorem 6.8, we can relate these integrals to derivatives, obtaining Cauchy's formulae. The special case $k = 0$ in 10.8 yields Liouville's theorem. See also Exercise 14.5.

## Multiplication of power series

Formally,

$$\sum_{n=0}^{\infty} a_n z^n \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} c_n z^n,$$

where $c_n = \sum_{r=0}^{n} a_r b_{n-r}$, the expression for $c_n$ being obtained by noting that terms in $z^n$ arise as products $a_r z^r \times b_s z^s$ where $r + s = n$. The formula is true whenever the series being multiplied converge absolutely. Most proofs of this result in the real case are highly technical. Fortunately Taylor's theorem yields a neat proof.

**14.10 Multiplication theorem for power series.** Suppose that

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad \text{and} \quad g(z) = \sum_{n=0}^{\infty} b_n z^n$$

are complex power series with radii of convergence $R_1$ and $R_2$, respectively. Let $h(z) = \sum_{n=0}^{\infty} c_n z^n$, where $c_n = \sum_{r=0}^{n} a_r b_{n-r}$. Then $\sum_{n=0}^{\infty} c_n z^n$ has radius of convergence at least $R := \min\{R_1, R_2\}$ and $h(z) = f(z)g(z)$ for $|z| < R$.

**Proof** In $\mathrm{D}(0; R)$ both $f$ and $g$ are holomorphic and we have $a_n = f^{(n)}(0)/n!$ and $b_n = g^{(n)}(0)/n!$. The product $fg$ is also holomorphic in $\mathrm{D}(0; R)$ and is represented there by a Taylor series

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

where

$$n! c_n = (fg)^{(n)}(0) = \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} f^{(r)}(0) g^{(n-r)}(0) = n! \sum_{r=0}^{n} a_r b_{n-r}.$$

Here we have used Leibniz' formula for the $n$th derivative of a product. (Those not familiar with this can check it for small values of $n$ by hand using the product rule for differentiation; for general $n$ it is proved by induction in just the same way that the binomial expansion for a positive integer exponent can be proved.) □

**14.11 Examples (products of power series).**

- The exponential series $\sum_{n=0}^{\infty} z^n/n!$ has infinite radius of convergence. The Multiplication theorem allows us to multiply the series for $e^{az}$ and $e^{bz}$.

Doing so and putting $z = 1$ gives, after some manipulation, $e^{a+b} = e^a e^b$ for all $a, b \in \mathbb{C}$ (cf. 7.3).

- The $n$th Hermite function $H_n$ is defined by

$$H_n(t) = (-1)^n e^{\frac{1}{2}t^2} \left(\frac{d}{dt}\right)^n e^{-t^2} \quad (n = 0, 1, 2, \ldots).$$

We claim that

$$\sum_{n=0}^{\infty} H_n(t) \frac{x^n}{n!} = e^{-\frac{1}{2}t^2 + 2xt - x^2} \quad (x, t \in \mathbb{R});$$

this gives a generating function for the Hermite functions. To prove this, note that $e^{-\frac{1}{2}t^2+2xt-x^2} = e^{\frac{1}{2}t^2} e^{-(x-t)^2}$. By the Chain rule, $e^{-z^2}$ is holomorphic in $\mathbb{C}$. It has a Taylor expansion

$$e^{-z^2} = \sum_{n=0}^{\infty} \left[\left(\frac{d}{dz}\right)^n e^{-z^2}\right]_{z=a} \frac{(z-a)^n}{n!},$$

for any $a \in \mathbb{C}$. Now put $z = x - t$ and $a = -t$ and the required formula drops out.

The Hermite functions are of importance in mathematical physics and elsewhere. Differentiation of the series $\sum H_n(t)x^n/n!$ with respect to either $x$ or $t$ can be proved to be legitimate. This enables the generating function to be used, painlessly, to obtain assorted recurrence relations for the functions $H_n$ and also the differential equation $H_n''(t) = (t^2 - 2n - 1)H_n(t)$.

## A primer on uniform convergence

Let $S$ be a non-empty subset of $\mathbb{C}$ and let $\{f_n(z)\}$ be a sequence of complex-valued functions which converges for each $z \in S$. How fast the sequence converges to its limit may vary from one point to another, as we illustrate in 14.12. When we need to consider some process, such as integration, which involves a *variable* point in $S$, we need some control over the rate of convergence if the limit function is to behave well. Uniform convergence gives such control.

**14.12 Introductory example.** Consider the sequence $\{f_n(z)\}$ in $S := D(0; 1)$, where $f_n(z) = z^n$. For each fixed $z \in S$ we certainly have $z^n \to 0$. In detail: $|z^n| < \varepsilon$ ($< 1$) if and only if $n \log |z| < \log \varepsilon$. Realizing that both sides are negative we see that this inequality holds when $n \geqslant N$, provided $N$ is chosen so that $N > |\log\varepsilon| / |\log|z||$. The critical observation now is that $N$ depends both on $\varepsilon$ and on $z$. There is no $N$ such that $|z^n| < \varepsilon$ for all $n \geqslant N$, simultaneously for all $z \in \mathrm{D}(0;1)$: the closer $|z|$ is to 1, the larger we have to take $n$ to bring $z^n$ within a distance $\varepsilon$ of the limiting value 0.

Now consider the same sequence, but in $S := \overline{\mathrm{D}}(0;1/2)$. Here $|z| < 1/2$ implies that $1/|\log|z|| < 1/|\log 1/2|$. Choose a natural number $N$ with $N > |\log\varepsilon| / |\log 1/2|$. Then $|z^n| < \varepsilon$ whenever $n \geqslant N$ *for all* $z \in S$ *at the same time.* There is nothing special about the choice of 1/2 as the radius of the disc, save that it must be *strictly* less than 1: for any $\delta$ with $0 < \delta < 1$,

$$(\forall\varepsilon > 0)(\exists N)\ \big(n \geqslant N \implies |z^n| < \varepsilon \text{ whenever } |z| \leqslant 1-\delta\big).$$

We stress that $N$ here can be chosen to be independent of $z$ (though it does, of course, depend on both $\delta$and $\varepsilon$).

In summary, what is happening here is that the rate of convergence of $\{z^n\}$ gets slower and slower as $|z|$ gets closer and closer to 1. But there is a minimum, worst case, rate of convergence so long as $|z|$ is restricted to some closed disc $\mathrm{D}(0;1-\delta)$ with $0 < \delta < 1$. Experimentation with a calculator may help you to grasp the point at issue.

**14.13 Uniform convergence of sequences.** Suppose that $\{f_n\}$ is a sequence of complex-valued functions defined on some set $S$. We say $\{f_n\}$ **converges pointwise** to $f$ on $S$ (and write $f_n \to f$) if, for each fixed $z \in S$, the complex sequence $\{f_n(z)\}$ converges to $f(z)$. We say $\{f_n\}$ **converges uniformly on** $S$ to $f$ (and write $f_n \xrightarrow{\mathrm{u}} f$ on $S$) if

$$\alpha_n := \sup\{\,|f_n(z) - f(z)| \,:\, z \in S\,\} \to 0 \text{ as } n \to \infty.$$

This is a convenient way of expressing, and later for working with, the condition

$$(\forall\varepsilon > 0)(\exists N_\varepsilon)\ (n \geqslant N_\varepsilon \implies (\forall z \in S)\ |f_n(z) - f(z)| < \varepsilon).$$

On the other hand, $f_n \to f$ pointwise on $S$ if

$$(\forall z \in S)\ (\forall\varepsilon > 0)\ (\exists N_\varepsilon(z))\ (n \geqslant N_\varepsilon(z) \implies |f_n(z) - f(z)| < \varepsilon);$$

here $N_\varepsilon(z)$ may depend on $z$. Obviously, $f_n \xrightarrow{\mathrm{u}} f$ on $S$ implies that $f_n \to f$ on $S$. The key difference between the two modes of convergence is this: in uniform convergence there is a single $N_\varepsilon$ which serves as $N_\varepsilon(z)$ for all $z$.

**14.14 Examples (uniform and non-uniform convergence).**

- For the sequence $\{f_n(z)\}$ in 14.12 we have $f_n(z) = z^n$ and $f(z) = 0$, so

$$\text{for } S = \mathrm{D}(0;1): \qquad \alpha_n := \sup\{|z^n - 0| \,:\, |z| < 1\} = 1 \not\to 0 \text{ as } n \to \infty;$$

whereas

$$\text{for } S = \overline{\mathrm{D}}(0;1-\delta): \quad \alpha_n := \sup\{\,|z^n - 0| \,:\, |z| \leqslant 1-\delta\,\} = (1-\delta)^n \to 0 \text{ as } n \to \infty.$$

Convergence is not uniform on the whole of $\mathrm{D}(0;1)$ but is uniform over any closed subdisc $\overline{\mathrm{D}}(0;1-\delta)$.

- Let $f_n(z) = (1+n^2z^2)^{-1}$ on $\mathrm{D}(0;1)$. Here, for fixed $z$,

$$f_n(z) \to f(z) = \begin{cases} 0 & \text{if } z \neq 0, \\ 1 & \text{if } z = 0. \end{cases}$$

Therefore, noting that $z = 0$ can be omitted when calculating the supremum because $f_n(0) - f(0) = 0$, we have

$$\alpha_n := \sup\{\big|(1+n^2z^2)^{-1}\big| \,:\, 0 < |z| < 1\} = 1 \not\to 0 \text{ as } n \to \infty.$$

So convergence is not uniform.

The latter example exhibits an interesting but disquieting feature: each $f_n$ is continuous on $\mathrm{D}(0;1)$, yet the limit function $f$ fails to be continuous at 0. Put another way,

$$\lim_{z\to 0}\Big\{\lim_{n\to\infty} f_n(z)\Big\} = 0 \quad \text{but} \quad \lim_{n\to\infty}\Big\{\lim_{z\to 0} f_n(z)\Big\} = 1.$$

(So we have an instance of limit processes not commuting with one another.)

The next result shows that this phenomenon cannot occur for uniformly convergent sequences.

**14.15 Limits and continuity.** Let $\{f_n\}$ be a sequence of continuous functions on a set $S$ and assume $f_n \xrightarrow{\text{u}} f$ on $S$. We claim that $f$ is continuous. The proof is a classic piece of $\varepsilon$-ology. We fix the obligatory $\varepsilon > 0$ and let $z \in S$. By uniform convergence we can find $N$ such that

$$n \geqslant N \implies (\forall w \in S)\ |f_n(w) - f(w)| < \varepsilon.$$

By continuity of $f_N$ at $z$ there exists $\delta > 0$ such that

$$|z - w| < \delta \implies |f_N(z) - f_N(w)| < \varepsilon$$

($\delta$ depends on $N$, but $N$ is fixed). For $|z - w| < \delta$,

$$\begin{aligned}|f(z) - f(w)| &= |f(z) - f_N(z) + f_N(z) - f_N(w) + f_N(w) - f(w)| \\ &\leqslant |f(z) - f_N(z)| + |f_N(z) - f_N(w)| + |f_N(w) - f(w)| < 3\varepsilon.\end{aligned}$$

This suffices to prove our claim.

**14.16 Uniformly convergent series.** As always, we transfer terminology and results about sequences across to the setting of series by taking sequences of partial sums. Given $\sum_{k=0}^{\infty} u_k$, let $f_n = u_0 + \cdots + u_n$. We say that $\sum_{k=0}^{\infty} u_k$ **converges uniformly on a set** $S$ if $\{f_n\}$ converges uniformly on $S$.

It is seldom easy to work out either the pointwise limit

$$f(z) := \lim_{n \to \infty} f_n(z) = \sum_{k=0}^{\infty} u_k(z)$$

or $\alpha_n := \sup\{\, |f_n(z) - f(z)| \ : \ z \in S\}$, the test sequence for uniform convergence. Fortunately there is a user-friendly sufficient condition for uniform convergence of a series. It is not a necessary condition.

**14.17 Weierstrass' M-test.** The series $\sum u_k$ converges uniformly on $S$ if there exist real numbers $M_k$ such that

$$(\forall k) \ \ |u_k(z)| \leqslant M_k \text{ for all } z \in S \quad \text{and} \quad \sum M_k \text{ converges.}$$

**Proof** We invoke the Cauchy convergence principle (3.21). This states that a sequence $\{z_n\}$ of complex numbers converges if and only if

$$(\forall \varepsilon > 0)(\exists N) \ (m, n \geqslant N \implies |z_m - z_n| < \varepsilon).$$

Let $f_n := u_0 + \cdots + u_n$. For each $z \in S$ and $n > m$,

$$|f_m(z) - f_n(z)| = |u_{m+1}(z) + \cdots + u_n(z)| \leqslant M_{m+1} + \cdots + M_n \to 0 \text{ as } m, n \to \infty$$

by the Cauchy condition applied to the partial sums of the series $\sum M_n$. Hence $\{f_n(z)\}$ satisfies the Cauchy condition and so converges, to $f(z)$, say. Thus the series $\sum u_n$ converges pointwise. To check that convergence is uniform, take the limit as $m \to \infty$ in the displayed line (with $z$ fixed) to get

$$(\forall z \in S)\ \ |f(z) - f_n(z)| \leqslant \sum_{k=n+1}^{\infty} M_k \to 0 \text{ as } n \to \infty. \qquad \square$$

**14.18 Power series and uniform convergence.** Let $\sum c_k z^k$ be a complex power series with radius of convergence $R > 0$. Then the series converges uniformly on $\overline{\mathrm{D}}(0; R-\delta)$ for any $\delta > 0$ (or on any closed disc $\overline{\mathrm{D}}(0; S)$ if $R = \infty$). To prove this, note that

$$\left|c_k z^k\right| \leqslant M_k := |c_k|\,(R-\delta)^k \quad \text{for } |z| \leqslant (R-\delta).$$

By the Radius of convergence lemma (6.6), $\sum M_k$ converges, so Weierstrass' M-test applies. Likewise, $\sum c_k z^k$ converges uniformly on any circle $\gamma(0; r)^*$ for which $0 \leqslant r < R$.

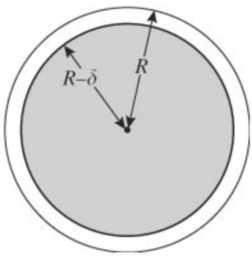



**Figure 14.1** Pointwise convergence vs. uniform convergence

In general $\sum c_k z^k$ does *not* converge uniformly on $\mathrm{D}(0; R)$. Consider, for example, the geometric series $\sum_{k=0}^{\infty} z^k$. In this case $R = 1$ and, for $|z| < 1$,

$$f_n(z) := 1 + z + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z},$$

which converges pointwise to $f(z) = (1 - z)^{-1}$. Then

$$\begin{aligned} \alpha_n := \sup\{\, |f_n(z) - f(z)| \ :\ |z| < 1 \,\} &= \sup\Big\{ \left|\frac{z^{n+1}}{1-z}\right| \ :\ |z| < 1 \Big\} \\ &\geqslant \sup\Big\{ \frac{x^{n+1}}{1-x} \ :\ 0 \leqslant x < 1 \Big\}. \end{aligned}$$

By differentiating $x^{n+1}/(1-x)$ we find that this function is strictly increasing on $[0,1)$. It tends to infinity as $x$ increases towards 1. Therefore $\alpha_n$ is not finite for any $n$, and we certainly do not have $\alpha_n \to 0$. Therefore $\sum_{k=0}^{\infty} z^k$ does *not* converge uniformly on the whole of $\mathrm{D}(0;1)$.

Finally, we prove that limits and integrals can be interchanged when convergence is uniform.

**14.19 Interchange theorem for uniformly convergent sequences and series.** Let $\gamma$ be a path with parameter interval $[\alpha, \beta]$.

(1) Let $\{f_n\}$ be a sequence of continuous complex-valued functions which converges uniformly on $\gamma^*$ to a (necessarily continuous) function $f$. Then

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

(2) Let $\{u_k\}$ be a sequence of continuous complex-valued functions such that $\sum u_k(z)$ converges uniformly on $\gamma^*$. Then

$$\sum_{k=0}^{\infty} \int_\gamma u_k(z)\,\mathrm{d}z = \int_\gamma \sum_{k=0}^{\infty} u_k(z)\,\mathrm{d}z.$$

**Proof** It suffices to prove (1). By the Estimation theorem,

$$\begin{aligned}\left|\int_\gamma f_n(z)\,\mathrm{d}z - \int_\gamma f(z)\,\mathrm{d}z\right| &= \left|\int_\gamma (f_n(z) - f(z))\,\mathrm{d}z\right| \\ &\leqslant \sup_{t\in[\alpha,\beta]} |f_n(\gamma(t)) - f(\gamma(t))| \times \mathrm{length}(\gamma).\end{aligned}$$

The final expression tends to 0 because the convergence is uniform. □

## Exercises

14.1 Find an expansion $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ valid in the disc $\mathrm{D}(a;r)$ when

$$\text{(i) } f(z) = \sin^2 z \quad (a=0), \qquad \text{(ii) } f(z) = (1+z)^{-1} \quad (a=\mathrm{i}),$$
$$\text{(iii) } f(z) = \mathrm{e}^z \quad (a=1).$$

(Hint: make use of known expansions.)

14.2 In the plane cut along $(-\infty, 0]$, define the square root of $z$ to be $f(z) = |z|^{1/2}\,\mathrm{e}^{\mathrm{i}\theta/2}$ $(0 \neq z = |z|\mathrm{e}^{\mathrm{i}\theta},\ -\pi < \theta \leqslant \pi)$. By using the identity

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

prove that $f'(z)$ exists and equals $f(z)/(2z)$ in $\mathbb{C} \setminus (-\infty, 0]$. By considering

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

for $|z-1| < 1$, obtain the Taylor expansion of $f$ in $\mathrm{D}(1;1)$.

14.3 Suppose that $f(z) = \sum_{n=0}^{\infty} c_n z^n$ for $z \in \mathbb{C}$. Prove that, for all $R$,

$$\sum_{n=0}^{\infty} |c_n|\, R^n \leqslant 2M(2R), \quad \text{where } M(r) := \sup\{\, |f(z)| \,:\, |z| = r\,\}.$$

14.4 Let $f$ be holomorphic in $\mathbb{C}$. Use 14.7 to prove that if $|f(z)| \leqslant M|z|^{\alpha}$, where $0 < \alpha < 1$, then $f$ is identically zero in $\mathbb{C}$.

14.5 Suppose that $f(z) = \sum_{n=0}^{\infty} c_n z^n$ for $z \in \mathrm{D}(0;R)$.

(i) Prove that

$$\frac{1}{2\pi}\int_0^{2\pi} \left|f(r\mathrm{e}^{\mathrm{i}\theta})\right|^2 \mathrm{d}\theta = \sum_{n=0}^{\infty} |c_n|^2\, r^{2n} \quad (0 \leqslant r < R).$$

(Hint: use the fact that $|f(z)|^2 = f(z)\overline{f(z)}$ and justify two interchanges of summation and integration. For the first of these justifications you will need the fact that $f$ is bounded on $\gamma(0;r)^*$.)

(ii) Suppose that $f \in \mathrm{H}(\mathbb{C})$ and that $f$ is bounded. Use (i) to deduce that $f$ is constant.

14.6 Let $f$ have a power series expansion $f(z) = \sum_{n=0}^{\infty} c_n z^n$ in $\mathrm{D}(0;R)$. Use the result of Exercise 14.5 to prove that, for $r < R$,

$$\int_0^{2\pi} |f(r\mathrm{e}^{\mathrm{i}\theta}) - P(r\mathrm{e}^{\mathrm{i}\theta})|^2\, \mathrm{d}\theta$$

attains its minimum over all polynomials $p$ of degree $k$ when $p(z) = \sum_{n=0}^{k} c_n z^n$.

The remaining exercises assume familiarity with uniform convergence.

14.7 Let $G$ be an open set and suppose that $\{f_k\}$ is a sequence of functions such that $f_k \in \mathrm{H}(G)$ $(k = 1, 2, \ldots)$ and $f_k \to f$ uniformly on $G$.

(i) Use Morera's theorem to prove that $f \in \mathrm{H}(G)$.

(ii) Use Cauchy's formula for derivatives to prove that

$$\lim_{k\to\infty} f_k^{(n)}(a) = f^{(n)}(a) \quad (n = 1, 2, \ldots).$$

Deduce corresponding results for uniformly convergent series of holomorphic functions.

14.8 Prove that, for each $\delta > 0$, the series $\sum_{n=0}^{\infty} n^{-z}$ converges uniformly on $\{ z : \operatorname{Re} z > 1+\delta \}$. Deduce that the series defines a holomorphic function $\zeta(z)$ in $\{ z : \operatorname{Re} z > 1 \}$. (This is the **Riemann zeta function**, of great interest in number theory. It is discussed briefly in the Appendix.)

14.9 By proving that the series converges uniformly on any disc $\mathrm{D}(a; r)$ containing no integer, prove that $\sum_{n=-\infty}^{\infty} (z-n)^{-2}$ defines a function holomorphic in $\mathbb{C} \setminus \mathbb{Z}$.

(In neither of the last two exercises does the given series converge uniformly on the whole of the region of holomorphy.)

# 15 Zeros of holomorphic functions

Let $f$ be holomorphic in an open set $G$. In this chapter we investigate the set

$$Z(f) := \{ z \in G : f(z) = 0 \}$$

of points at which $f$ takes the value $0$. There are two reasons for doing this. The first is that $1/f$ fails to be defined at $a$ if $f(a) = 0$. Zeros in the denominators of rational functions and of functions like $\cot z = \cos z / \sin z$ give rise to singularities (informally, points where an otherwise holomorphic function fails to be holomorphic). Integrating a function round a contour inside which it has one or more singularities will in general give a non-zero result (recall 10.4 and the Cauchy formulae). Understanding, and exploiting, singularities will be the thrust of Chapters 17–22.

Our second reason for investigating zeros is more theoretical. Taylor's theorem implies that holomorphic functions are locally representable by power series, with the coefficients expressible in terms of the derivatives. The consequences of this fact are surprising and far-reaching. It turns out that a holomorphic function in a *region* cannot be zero except at isolated points without being identically zero (the Identity theorem (15.8)).

## Characterizing zeros

**15.1 Definitions (zeros and their orders).** Suppose that $f$ is holomorphic at $a$, that is, suppose that $f \in \mathrm{H}(\mathrm{D}(a; r))$ for some $r > 0$. The point $a$ is said to be a **zero** of $f$ if $f(a) = 0$. We say that the zero $a$ of $f$ is of **order** $m$ if

$$0 = f(a) = f'(a) = \cdots = f^{(m-1)}(a) \quad \text{and} \quad f^{(m)}(a) \neq 0.$$

Zeros of orders $1, 2, \ldots$ are called **simple, double, . . .** . For convenience, we adopt the convention that $f$ has a zero of order $0$ at $a$ if $f$ is holomorphic at $a$ and $f(a) \neq 0$.

**15.2 Examples (orders of zeros).** Here are some simple examples. Recall the results in 7.10 on the location of zeros of trigonometric and hyperbolic functions and of functions related to exponentials.

- $(z-a)^m$ has a zero of order $m$ at $a$.
- $\sin z = 0$ if and only if $z = k\pi$ $(k \in \mathbb{Z})$. At $z = k\pi$, $(\mathrm{d}/\mathrm{d}z)\sin z = \cos z \neq 0$. Hence all the zeros of $\sin z$ are simple. Likewise, all the zeros of $\cos z$, $\sinh z$, and $\cosh z$ are simple.
- $1+\mathrm{e}^z = 0$ if and only if $z = (2k+1)\pi\mathrm{i}$ $(k \in \mathbb{Z})$. By 7.3(3), $(\mathrm{d}/\mathrm{d}z)(1+\mathrm{e}^z) = \mathrm{e}^z \neq 0$, so all the zeros of $1+\mathrm{e}^z$ are simple.

To handle functions more complicated than those above we need further techniques. See 15.4 and the examples in 15.5.

The proof of part of the next result is rather technical but the result is extremely useful.

**15.3 Characterization theorem for zeros of order $m$.** Let $f \in \mathrm{H}(\mathrm{D}(a;R))$ and suppose that $f$ has Taylor expansion $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ in $\mathrm{D}(a;R)$. Then the following are equivalent:

(1) $0 = f(a) = f'(a) = \cdots = f^{(m-1)}(a)$ and $f^{(m)}(a) \neq 0$;

(2) $f(z) = \sum_{n=m}^{\infty} c_n(z-a)^n$, where $c_m \neq 0$;

(3) $f(z) = (z-a)^m g(z)$, where $g \in \mathrm{H}(\mathrm{D}(a;R))$ and $g(a) \neq 0$;

(4) there exists a non-zero constant $C \in \mathbb{C}$ such that

$$\lim_{z\to a}(z-a)^m f(z) \text{ exists and equals } C.$$

**Proof** The equivalence of (1) and (2) comes from the facts collected together in 14.4.

Assume (2). Then we let

$$g(z) := \sum_{n=m}^{\infty} c_n(z-a)^{n-m} = \sum_{k=0}^{\infty} c_{m+k}(z-a)^k.$$

Since $g$ is defined by a convergent power series, it is holomorphic in $\mathrm{D}(a;R)$. Also, $g(a) = c_m \neq 0$. Hence (4) holds. The continuity of $g$ and the algebra of limits show that (4) implies (3).

Finally, assume that (3) holds. Given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left|(w-a)^{-m}f(w) - C\right| < \varepsilon \quad \text{whenever } 0 < |w-a| < \delta.$$

Take $r < \min\{\delta, R\}$. Then

$$|w-a| = r \Longrightarrow \left|(w-a)^{-m} f(w)\right| \leqslant |C| + \varepsilon \qquad \text{(by 1.9(2))}$$
$$\Longrightarrow |f(w)| \leqslant (|C| + \varepsilon) r^m.$$

The estimate for $c_n$ given in 14.7 gives

$$|c_n| \leqslant (|C| + \varepsilon) r^{m-n}.$$

If $n < m$ then we can make $r^{m-n}$ arbitrarily small by taking $r$ sufficiently small. The constant $c_n$ is independent of $r$ and so must be zero. We now have $f(z) = \sum_{n=m}^{\infty} c_n (z-a)^n$. As in the proof above we have

$$c_m = \lim_{z \to a} (z-a)^{-m} f(z) = C \neq 0. \qquad \square$$

**15.4 Compound zeros.** It could be shown directly from the definition in 15.1 that $z^2 \sin z$ has a zero of order 3 at $z = 0$. But calculations of derivatives of compound functions can be messy. It is much easier to use instead the fact that if $f$ and $g$ have zeros of order $m \geqslant 0$ and $n \geqslant 0$, respectively, at $z = a$ then $fg$ has a zero of order $m + n$ at $a$. This follows immediately from 15.3(4) and the algebra of limits.

**15.5 Examples (compound zeros).**

- $z^2 \sin^4 z$ has a zero of order $2 + 4 = 6$ at $z = 0$ and zeros of order $0 + 4 = 4$ at $z = k\pi$ $(k \in \mathbb{Z} \setminus \{0\})$.
- $\cosh^3 z$ has triple zeros at $z = (2k+1)\pi i/2$ $(k \in \mathbb{Z})$, since all the zeros of $\cosh z$ are simple.

## The Identity theorem and the Uniqueness theorem

Our analysis of zeros so far has focused on a single point $a$. Now we shall consider the set $Z(f)$ of *all* zeros of a function $f$ holomorphic in a region $G$. (Recall that a region is by definition a non-empty open connected set.) We shall see that clustering of zeros at a limit point in $G$ is sufficient to force the function to be identically zero. We begin with a recap on limit points. We defer examples until we have our two major theorems in place.

**15.6 Limit points revisited.** The definition was given in 3.6(2): a point $a \in \mathbb{C}$ is a **limit point** of a set $S$ if, for every $\varepsilon > 0$, $\mathrm{D}'(a;\varepsilon) \cap S \neq \varnothing$, where $\mathrm{D}'(a;\varepsilon)$ is the punctured disc $\{ z \in \mathbb{C} : 0 < |z-a| < \varepsilon \}$.

Let $G$ be an open set and let $S \subseteq G$. For our purposes here, the following are important prototypical examples of the occurrence of at least one limit point of $S$ in $G$:

- $S = \{z_n\}_{n \geqslant 1}$, with the points $z_n$ distinct and $z_n \to z \in G$ ($z$ is a limit point of $S$);
- $S$ is a segment $[\alpha, \beta]$, with $\alpha \neq \beta$ (every point of $S$ is a limit point of $S$);
- $S$ is an open disc $\mathrm{D}(a;r)$ (the set of limit points of $S$ in $G$ is $\overline{\mathrm{D}}(a;r) \cap G$, and this contains $\mathrm{D}(a;r)$).

**15.7 Identity theorem (the special case in which the region is a disc).** Suppose that $f \in \mathrm{H}(\mathrm{D}(a;r))$ and that $f(a) = 0$. Then either

(1) $f$ is identically zero in $\mathrm{D}(a;r)$, or

(2) the zero of $f$ at $a$ is **isolated**, that is, there exists $\varepsilon > 0$ such that the punctured disc $\mathrm{D}'(a;\varepsilon)$ contains no zeros of $f$.

Consequently, if $a$ is a limit point of $Z(f)$ then $f \equiv 0$ in $\mathrm{D}(a;r)$.

**Proof** By Taylor's theorem (14.4),

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n \quad (z \in \mathrm{D}(a;r)).$$

There are two possibilities. If all $c_n = 0$ then (1) holds. Otherwise there exists a smallest integer $m > 0$ such that $c_m \neq 0$ and we may write

$$f(z) = (z-a)^m g(z), \qquad \text{where } g(z) := \sum_{k=0}^{\infty} c_{k+m} (z-a)^k.$$

The series defining $g$ has radius of convergence at least $r$. Hence, by Theorem 6.8(2), $g$ is continuous on $\mathrm{D}(a;r)$. Because $g(a) = c_m \neq 0$ and $g$ is continuous at $a$, there is some $\varepsilon > 0$ such that $g(z) \neq 0$ in $\mathrm{D}'(a;\varepsilon)$. Throughout this punctured disc we also have $f(z) \neq 0$. □

The following simple example shows that we certainly cannot extend the Identity theorem to an arbitrary open set $G$. Let

$$f(z) := \begin{cases} 1 & \text{if } z \in \mathrm{D}(-2;1), \\ 0 & \text{if } z \in \mathrm{D}(2;1). \end{cases}$$

Then $f$ is not identically zero in $G := \mathrm{D}(-2;1) \cup \mathrm{D}(2;1)$ despite every point of $\mathrm{D}(2;1)$ being a limit point of $Z(f)$. Of course, $G$ here is not connected.

**15.8 Identity theorem (general form).** Let $G$ be a region and suppose that $f \in \mathrm{H}(G)$. Assume that the set $Z(f)$ of zeros of $f$ has a limit point in $G$. Then $f$ is identically zero in $G$.

In particular, if $f \equiv 0$ on some open disc $\mathrm{D}(a;r) \subseteq G$, then $f$ is identically zero in $G$.

**Proof** The proof is topological in nature. Our strategy is to prove that $E$, the set of limit points of $Z(f)$ in $G$, is such that

(i) $Z(f) \supseteq E$, and

(ii) $E$ and $G \setminus E$ are both open.

Since $G$ is a region and $E \neq \varnothing$ by assumption, (ii) implies that $E = G$. Thence, by (i), $f \equiv 0$ in $G$.

To prove (i), let $a \in E$. For each $n$ there exists $a_n \in \mathrm{D}'(a;1/n)$ such that $f(a_n) = 0$. By continuity of $f$ we have $f(a) = 0$, so $a \in Z(f)$.

To prove (ii), first let $a \in E$. Then 15.7 implies that $f \equiv 0$ in $\mathrm{D}(a;r)$ for some $r > 0$. But then $\mathrm{D}(a;r) \subseteq E$, so $E$ is open. To show $G \setminus E$ is open, take $a \in G \setminus E$. Since $a$ is not a limit point of $Z(f)$, there exists $\mathrm{D}'(a;r)$ in which $f(z)$ is never zero. No point of $\mathrm{D}(a;r)$ belongs to $E$, so $\mathrm{D}(a;r) \subseteq G \setminus E$. □

The Identity theorem leads to a very useful uniqueness theorem via the simple observation that $f(z) = g(z)$ if and only if $(f-g)(z) = 0$, for functions $f$ and $g$. The theorem is an immediate consequence of 15.8.

**15.9 Uniqueness theorem.** Suppose that $G$ is a region, that $f$ and $g$ belong to $\mathrm{H}(G)$, and that $f(z) = g(z)$ for all $z \in S$, where $S$ has a limit point in $G$. Then $f \equiv g$ in $G$.

**15.10 Examples (Identity theorem and Uniqueness theorem).**

- Suppose that $f \in \mathrm{H}(\mathrm{D}(0;1))$ is such that $f(z) = 0$ whenever $z \in (0,1)$. Then $f \equiv 0$ in $\mathrm{D}(0;1)$ (since every point of $(0,1)$ is a limit point of $Z(f)$ in $\mathrm{D}(0;1)$).
- Suppose that $f \in \mathrm{H}(\mathbb{C})$ and that $f(1/n) = \sin(1/n)$ $(n = 1, 2, \ldots)$. Then 0 is a limit point of $S = \{\, 1/n : n = 1, 2, \ldots \}$ and $f(z) = \sin z$ on $S$. By the Uniqueness theorem, $f(z) = \sin z$ in $\mathbb{C}$.
- Suppose $f$ is holomorphic in $\mathbb{C} \setminus \{0\}$ and $f(z) = \sin(1/z)$ whenever $z = 1/(n\pi)$ $(n = 1, 2, \ldots)$. It does not follow that $f(z) = \sin(1/z)$ in $\mathbb{C} \setminus \{0\}$.

Indeed, $f \equiv 0$ would fit the given conditions. Here 0 is a limit point of $Z(g)$, where $g(z) := f(z) - \sin(1/z)$. However the Identity theorem does not apply because the limit point 0 is not in $\mathbb{C} \setminus \{0\}$, the region of holomorphy of $g$.

- We shall show that there is no $f \in \mathrm{H}(\mathrm{D}(0;1))$ such that $f(x) = |x|^3$ for $-1 < x < 1$. Certainly $\mathrm{D}(0;1)$ is a region and 0 is a limit point in $\mathrm{D}(0;1)$ of each of the segments $[0,1)$ and $(-1,0]$. Suppose, for a contradiction, that $f$ did exist. Then $f(z) = z^3$ on the segment $[0,1)$ in $\mathrm{D}(0;1)$. By the Uniqueness theorem, $f(z) = z^3$ throughout $\mathrm{D}(0;1)$. In the same way, $f(z) = -z^3$ on $(-1,0] \subseteq \mathrm{D}(0;1)$, so we have the required contradiction.

**15.11 Preservation of functional identities.** The Uniqueness theorem allows us to extend the domain of validity of certain functional identities, a procedure we alluded to in 7.7. The method is best illustrated by examples.

- The identity $\cos^2 z + \sin^2 z = 1$ holds when $z$ is real. Now 15.9 implies that it holds for all complex $z$, since 1 and $\cos^2 z + \sin^2 z$ are holomorphic in $\mathbb{C}$ and the real axis has limit points in $\mathbb{C}$ (every point of $\mathbb{R}$ is a limit point).
- Our second example concerns the binomial expansion. Given a negative integer $n$, let

$$f(z) = (1+z)^n \quad \text{and} \quad g(z) = \sum_{k=0}^{\infty} \frac{n(n-1)\dots(n-k+1)}{k!} z^k.$$

  Clearly $f$ is holomorphic except at $-1$. The series defining $g$ has radius of convergence 1, so $g \in \mathrm{H}(\mathrm{D}(0;1))$, by 6.8. It is well known that $f(x) = g(x)$ when $x$ is real and $|x| < 1$. The Identity theorem tells us that the equality holds throughout the region $\mathrm{D}(0;1)$. The requirement that $n$ be an integer can be relaxed though care is needed because non-integer powers produce multifunctions; see Exercise 14.2.

**15.12 Analytic continuation.** The Uniqueness theorem is the gateway to an important technique in complex function theory, known as analytic continuation. The objective is to extend a given function $f$, holomorphic in some region $G$, to a function $g \in \mathrm{H}(G')$, where $G'$ is a region strictly containing $G$. If such a function $g$ exists, then it is necessarily unique.

The idea is very simply illustrated. Let

$$f(z) = \sum_{n=0}^{\infty} z^n \qquad \text{and} \qquad g(z) = (1-z)^{-1}.$$

Here $f(z)$ is defined (and $f$ holomorphic) only in $\mathrm{D}(0;1)$, whereas $g$ is holomorphic in a much bigger region, $G' := \mathbb{C} \setminus \{1\}$; and $f = g$ in the common domain $\mathrm{D}(0;1)$. We say $g$ is a **direct analytic continuation** of $f$. We cannot hope to do any better in this case. Since $f$ is unbounded, the Boundedness theorem (3.24) tells us that $f$ cannot be extended to a function holomorphic in any open set containing $\overline{\mathrm{D}}(0;1)$.

In general, it will not be obvious whether we can extend a given holomorphic function. However there is a natural way to attempt this. Consider, for example, a function $f$ with $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ in $\mathrm{D}(a;R)$. The (unique) Taylor expansion of $f$ about $b \in \mathrm{D}'(a;r)$ is given by

$$f_1(z) := \sum_{n=0}^{\infty} \frac{f^{(n)}(b)}{n!}(z-b)^n;$$

this certainly converges in $\mathrm{D}(b;R-|b-a|)$, the largest open disc centre $b$ contained in $\mathrm{D}(a;R)$. *But it may converge in a larger disc, $D$ say.* If so, we may extend $f$ to $g \in \mathrm{H}(\mathrm{D}(a;R) \cup D)$ by taking

$$g(z) = \begin{cases} f(z) & \text{for } z \in \mathrm{D}(a;R), \\ f_1(z) & \text{for } z \in D; \end{cases}$$

see Fig. 15.1. We may then repeat the process, following a chain of overlapping discs.

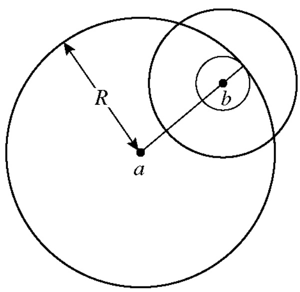



**Figure 15.1** Analytic continuation

For example, $\sum_{n=0}^{\infty} 2^n(z-\mathrm{i})^n/(1-2\mathrm{i})^{n+1}$ converges for $|z-\mathrm{i}| < |1-2\mathrm{i}| = \sqrt{5}$ and provides the analytic continuation of $\sum_{n=0}^{\infty}(z/2)^n$ from $\mathrm{D}(0;2)$ into the disc $\mathrm{D}(\mathrm{i};\sqrt{5})$. We can then re-expand about any point of $\mathrm{D}(\mathrm{i};\sqrt{5})$, for example the point $2+\mathrm{i}$ (which lies outside the original disc $\mathrm{D}(0;2)$), and so on.

It can be proved that if analytic continuation is possible at all then it can be accomplished by following a chain of overlapping discs along some path $\gamma$.

[Continuation along paths links up with the Riemann surfaces approach to multifunctions and with simple connectedness. In a simply connected region, the extension is independent of the path. In other cases it may be possible to obtain an extension by following a multibranch, returning to the starting point with a different function value.]

Analytic continuation plays a very important part in the theory of some famous functions of a complex variable. In particular, it is of central importance in the exploration of the properties of the Riemann zeta function (introduced fleetingly in Exercise 14.7) and hence in the derivation of the Prime number theorem; see the Appendix.

## Counting zeros

The theorems in this section have close affinities with the results on argument in Chapters 9 and 11. Exercise 15.16 (advanced track) pursues these connections. Here we adopt a less sophisticated, basic track approach. Theorem 15.13 can be derived from Cauchy's residue theorem (18.3) once we have that result available. At this stage, we give a direct proof, imitating the argument given in 18.3. See also Exercise 18.9, which extends Theorem 15.13.

**15.13 Theorem (counting zeros).** Let $f$ be holomorphic inside and on a positively oriented contour $\gamma$. Let $f$ be non-zero on $\gamma$ and have $N$ zeros inside $\gamma$. Then

$$\frac{1}{2\pi \mathrm{i}}\int_\gamma \frac{f'(z)}{f(z)}\,\mathrm{d}z = N.$$

(A zero of order $m$ is counted $m$ times.)

**Proof** The function $f'/f$ is holomorphic inside and on $\gamma$ except at the zeros of $f$ lying inside $\gamma$. Suppose that these zeros are at $a_1, \ldots, a_N$, and are of orders $m_1, \ldots, m_N$.

We can find disjoint open discs $\mathrm{D}(a_k; r_k)$ $(k = 1, \ldots, N)$ such that there exists a function $g_k$ which is holomorphic and non-zero in $\mathrm{D}(a_k; r_k)$ and such that

$$f(z) = (z - a_k)^{m_k} g_k(z) \qquad (z \in \mathrm{D}(a_k; r_k))$$

(see 15.3). Then

$$\frac{f'(z)}{f(z)} = \frac{m_k}{z - a_k} + \frac{g_k'(z)}{g_k(z)} \qquad (z \in \mathrm{D}'(a_k; r_k)).$$

Define

$$F(z) := \begin{cases} \dfrac{f'(z)}{f(z)} - \displaystyle\sum_{j=1}^{N} \dfrac{m_j}{z-a_j} & \text{if } z \notin \displaystyle\bigcup_{j=1}^{N} \mathrm{D}(a_j; r_j), \\ \dfrac{g_k'(z)}{g_k(z)} - \displaystyle\sum_{\substack{1 \leqslant j \leqslant N, \\ j \neq k}} \dfrac{m_j}{z-a_j} & \text{if } z \in \mathrm{D}(a_k; r_k) \quad (k = 1, \ldots, N). \end{cases}$$

Then $F$ is holomorphic inside and on $\gamma$. By Cauchy's theorem, $\int_\gamma G(z)\,\mathrm{d}z = 0$. The required result now follows from the Fundamental integral (10.4). □

**15.14 Rouché's theorem.** Let $f$ and $g$ be holomorphic inside and on a contour $\gamma$ and suppose that $|f(z)| > |g(z)|$ on $\gamma^*$. Then $f$ and $f+g$ have the same number of zeros inside $\gamma$.

(The number of zeros is finite. This is a consequence of the Bolzano–Weierstrass theorem (3.23) and the Identity theorem (15.8), applied with some care.)

**Proof** Let $t \in [0,1]$. Since $|f(z)| > |g(z)|$ on $\gamma^*$, we have $(f+tg)(z) \neq 0$ for any $z \in \gamma^*$ (by 1.9(3)). Assume, without loss of generality, that $\gamma$ is positively oriented, and define

$$\varphi(t) = \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{(f' + tg')(z)}{(f+tg)(z)}\,\mathrm{d}z.$$

By 15.13, $\varphi(t)$ is the number of zeros of $f+tg$ inside $\gamma$. The function $\varphi$ is integer-valued; if it is continuous, it must be constant (see 3.25). In this event, $\varphi(0)$, the number of zeros of $f$ inside $\gamma$, equals $\varphi(1)$, the number of zeros of $f+g$ inside $\gamma$.

It is possible to establish continuity of $\varphi$ by citing a general theorem about functions defined by integrals. Here, alternatively, is a direct proof. Fix $t$ and consider

$$\varphi(t) - \varphi(s) = \frac{t-s}{2\pi \mathrm{i}} \int_\gamma \frac{(g'f - f'g)(z)}{(f+tg)(z)(f+sg)(z)}\,\mathrm{d}z.$$

By 1.18, we can find positive constants $M$ and $m$ such that, for all $z \in \tilde{\gamma}$, $|(g'f - f'g)(z)| \leqslant M$, $|g(z)| \leqslant M$, and $|(f+tg)(z)| \geqslant m$. Then

$$\begin{aligned} |(f+sg)(z)| &\geqslant |(f+tg)(z)| - |s-t|\,|g(z)| \quad \text{(by 1.9(3))} \\ &\geqslant \frac{1}{2}m \quad \text{if } |s-t| \leqslant \frac{m}{2M}. \end{aligned}$$

Hence, for $|s-t|$ sufficiently small,

$$|\varphi(t) - \varphi(s)| \leqslant \frac{|t-s|\,M}{\pi m} \times \text{length}(\gamma) \quad \text{(by 10.10)}.$$

We conclude that $\varphi$ is continuous at $t$. $\square$

In Chapter 16 we shall employ Rouché's theorem in the proofs of some important theoretical results. Here we illustrate how the theorem can be used to locate zeros of particular functions.

**15.15 Example (locating zeros via Rouché's theorem).** We show that the function $2+z^2-\mathrm{e}^{\mathrm{i}z}$ has precisely one zero in the open upper half-plane. We take $f(z)=2+z^2$ and $g(z)=-\mathrm{e}^{\mathrm{i}z}$ and let $\gamma$ be the semicircular contour shown in Fig. 10.1, with $R>\sqrt{3}$. For $z\in[-R,R]$,

$$|f(z)|\geqslant 2>1=|g(z)|$$

and, for $z=R\mathrm{e}^{\mathrm{i}\theta}$ $(0\leqslant\theta\leqslant\pi)$,

$$|f(z)|\geqslant R^2-2>1\geqslant \mathrm{e}^{-R\sin\theta}=|g(z)|\,.$$

We deduce from Rouché's theorem that $f(z)+g(z)=2+z^2-\mathrm{e}^{\mathrm{i}z}$ has the same number of zeros in $\{\, z\,:\,\operatorname{Im} z>0,\ |z|<R\,\}$ $(R>\sqrt{3})$ as has $f(z)=2+z^2$, that is, just one. This proves our claim.

## Exercises

15.1 Determine the orders of the zeros of the functions in Exercise 7.11.

15.2 Let $f$ and $g$ be holomorphic in $\mathrm{D}(a;r)$ for some $r>0$ and assume that $f(a)=g(a)=0$. Use 15.3 to prove that

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

if the right-hand side exists. (This is a complex form of L'Hôpital's rule.) Hence evaluate

$$\text{(i)}\ \lim_{z\to \mathrm{i}}\frac{1+\mathrm{e}^{\pi z}}{1+z^2},\quad \text{(ii)}\ \lim_{z\to 0}(\cot z-z^{-1}),\quad \text{(iii)}\ \lim_{z\to 1}\frac{1-z}{1-z\mathrm{e}^{\lambda(1-z)}}.$$

15.3 Suppose that $\{z_n\}$ is a sequence of distinct points in $D(0;1)$ such that $z_n \to 0$ as $n \to \infty$. Decide whether the following statements are true for all choices of $\{z_n\}$.

(i) If $f$ is holomorphic in $D(0;1)$ and $f(z_n) = \sin z_n$ for all $n$, then $f(z) = \sin z$ for all $z \in D(0;1)$.

(ii) There exists $f \in D(0;1)$ such that $f(z_n) = n$ for all $n$.

(iii) There exists $f \in D(0;1)$ such that $f(z_n) = 0$ when $n$ is even and such that $f(z_n) = z_n$ when $n$ is odd.

(iv) There exists $f \in D(0;1)$ which is such that $f(z_n) = (-1)^n z_n$ for every $n$.

Justify your answers. How do your answers change if we assume that $z_n \to 1$ instead of $z_n \to 0$?

15.4 Give an example of a function $f$ which is holomorphic and not identically zero in $D(0;1)$ and such that the set of limit points of $Z(f)$ is $\{\pm 1, \pm \mathrm{i}\}$.

15.5 Suppose that $f$ is defined on $D(0;3)$ and takes the value $(-1)^n$ at the point $(-1)^n(1+\frac{1}{n})$ $(n = 1, 2, \dots)$. By considering the sets

$$S_- := \{\, -1 - \tfrac{1}{2m+1} \ : \ m \geqslant 1 \,\} \text{ and } S_+ := \{\, 1 + \tfrac{1}{2m} \ : \ m \geqslant 1 \,\},$$

show that $f$ cannot be holomorphic in $D(0;3)$.

15.6 Find all functions $f$ which are holomorphic in $D(0;1)$ and such that $f(1/n) = n^2 f(1/n)^3$ for $n = 2, 3, 4, \dots$.

15.7 Let $\{c_n\}$ be a sequence of complex numbers such that $\sum |c_n|$ converges and $\sum_{n=0}^{\infty} c_n k^{-n} = 0$ for $k = 1, 2, 3, \dots$. Prove that $c_n = 0$ for all $n$.

15.8 Let $f \in H(D(0;1))$.

(i) Deduce from Exercise 5.11 or Exercise 6.6 that $g$ defined by

$$g(z) = f(z) - \overline{f(-\overline{z})} \quad (z \in D(0;1))$$

is holomorphic in $D(0;1)$.

(ii) Now suppose that $f$ takes real values on the imaginary axis. Prove that, for $x + \mathrm{i}y \in D(0;1)$,

$$u(x,y) = u(-x,-y) \quad \text{and} \quad v(x,y) = -v(-x,y),$$

where $u$ and $v$ denote the real and imaginary parts of $f$.

15.9 Give alternative derivations of the logarithmic and binomial expansions in 14.6 and in Exercise 14.2 by assuming suitable real expansions and appealing to the Uniqueness theorem.

15.10 Prove that the equation $z^5 + 15z + 1 = 0$ has precisely four solutions in the annulus $\{ z : \frac{3}{2} < |z| < 2 \}$.

15.11 Find the number of zeros of each of the following functions in $D(0;1)$:

$$\text{(i) } z^5 - 3z + 1, \quad \text{(ii) } z^7 + 2z^5 + 2z^2 + 6, \quad \text{(iii) } \cos \pi z - 100z^{100}.$$

15.12 Suppose that $R > 0$ is given. Prove that, if $N$ is sufficiently large,

$$\sum_{n=0}^{N} z^n/n! \neq 0 \quad \text{for all } z \in D(0;R).$$

15.13 Prove that, for $n = 3, 4, \dots$, the polynomial $z^n + nz - 1$ has $n$ zeros inside the circle with centre at 0 and radius $1 + \sqrt{2/(n-1)}$.

15.14 Show that, for each $\lambda > 1$, the equation $z + e^{-z} = \lambda$ has precisely one zero in the open right half-plane, and that this zero is real.

15.15 Suppose that $f$ is holomorphic inside and on $\gamma(0;1)$, with Taylor expansion $\sum_{n=0}^{\infty} c_n z^n$. Given that $f$ has $m$ zeros inside $\gamma(0;1)$, prove that

$$\min\{ |f(z)| : |z| = 1 \} \leqslant |c_0| + |c_1| + \cdots + |c_m| .$$

15.16 [This advanced track exercise provides an alternative proof of Rouché's theorem (15.14), making use of the concept of index introduced in 12.8. Suppose that $f$, $g$, and $\gamma$ satisfy the conditions of Rouché's theorem. Define $F$ by $F(z) = \big(f(z) + g(z)\big)/f(z)$, and let $\Gamma$ be the path $F \circ \gamma$. Prove that $\Gamma^* \subseteq D(1;1)$ and hence show that $n(\Gamma, 0) = 0$. By applying Theorem 15.13 to $F$, deduce Rouché's theorem.]

# 16 Holomorphic functions: further theory

Here we present some additional results about holomorphic functions and about conformal maps. These are all important tools for more advanced theory. While this chapter is not designated as 'advanced track', it could be omitted by readers eager to reach the applications-oriented later chapters.

## The Maximum modulus theorem

**16.1 Local maximum modulus theorem.** Suppose that $f \in \mathrm{H}(\mathrm{D}(a;R))$ and that $|f(z)| \leqslant |f(a)|$ for all $z \in \mathrm{D}(a;R)$. Then $f$ is constant.

**Proof** Fix $r$ with $0 < r < R$. By Cauchy's integral formula (13.1),

$$\begin{aligned} f(a) &= \frac{1}{2\pi \mathrm{i}} \int_{\gamma(a;r)} \frac{f(z)}{z-a}\,\mathrm{d}z \\ &= \frac{1}{2\pi \mathrm{i}} \int_0^{2\pi} \frac{f(a+r\mathrm{e}^{\mathrm{i}\theta}) r\mathrm{i}\mathrm{e}^{\mathrm{i}\theta}}{r\mathrm{e}^{\mathrm{i}\theta}}\,\mathrm{d}\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(a+r\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta. \end{aligned}$$

From this and from the hypothesis of the theorem we see that

$$|f(a)| \leqslant \frac{1}{2\pi} \int_0^{2\pi} \left|f(a+r\mathrm{e}^{\mathrm{i}\theta})\right| \mathrm{d}\theta \leqslant |f(a)|\,.$$

Therefore

$$\int_0^{2\pi} \left(|f(a)| - \left|f(a+r\mathrm{e}^{\mathrm{i}\theta})\right|\right) \mathrm{d}\theta = 0.$$

Since the integrand is continuous and is non-negative, it must be identically zero if the integral is to vanish. This is true for every $r < R$. It follows that $|f|$ is constant in $\mathrm{D}(a;R)$. By 5.12, $f$ itself must be constant. $\square$

**16.2 Maximum modulus theorem.** Let $G$ be a bounded region and let $f$ be holomorphic in $G$ and continuous on the closure $\overline{G}$ of $G$. Then $|f|$ attains its maximum on the boundary $\partial G = \overline{G} \setminus G$.

**Proof** The set $\overline{G}$ is bounded and closed, so on $\overline{G}$ the continuous function $|f|$ is bounded and attains its supremum $M$ at some point of $\overline{G}$ (by 3.24). Now assume that $|f|$ does not attain the value $M$ on $\partial G$. Then $|f(a)| = M$ for some $a \in G$. Since $G$ is open, there exists $R > 0$ such that $\mathrm{D}(a; R) \subseteq G$. By 16.1, $f$ is constant on $\mathrm{D}(a; R)$. Hence, by the Identity theorem, $f$ is constant in $G$ (see 15.8). By continuity, $f$ is constant on $\overline{G}$, and so attains its supremum at every point of $\overline{G} = \partial G \cup G$, contrary to hypothesis. □

The following corollary of the Maximum modulus theorem is frequently useful in applying the theorem. See Exercise 16.6 for further deductions from the same hypotheses.

**16.3 Schwarz' lemma.** Suppose that $f$ is holomorphic in $\mathrm{D}(0; R)$, that $f(0) = 0$, and that $|f(z)| \leqslant M$ in $\overline{\mathrm{D}}(0; R)$. Then

$$|f(z)| \leqslant \frac{M}{R}|z| \quad (|z| \leqslant R).$$

If equality occurs for some $z$ with $|z| < R$, then there exists a real constant $\lambda$ such that $f(z) = Mz\mathrm{e}^{\mathrm{i}\lambda}/R$ for $z \in \mathrm{D}(0; R)$.

**Proof** Since $f(0) = 0$, there exists $g \in \mathrm{H}(\mathrm{D}(0; R))$ such that $zg(z) = f(z)$ for all $z \in \mathrm{D}(0; R)$ (see 15.3). On $|z| = r < R$,

$$|g(z)| \leqslant |f(z)|/r \leqslant M/r.$$

Applying the Maximum modulus theorem to $g$ in $G = \mathrm{D}(0; r)$, we obtain $|g(z)| \leqslant M/r$ for $|z| \leqslant r$. Now let $r \to R$ to get $|g(z)| \leqslant M/R$ for $|z| < R$. This gives the required bound on $|f(z)|$ for any $z \neq 0$, and the inequality holds for $z = 0$ too since $f(0) = 0$. We leave the proof of the final claim as an exercise. □

## Holomorphic mappings

In many applications of conformal mapping it is necessary to construct a conformal map $f: G \to \widehat{G}$ between regions $G$ and $\widehat{G}$, such that the inverse mapping $f^{-1}: \widehat{G} \to G$ exists and is also conformal. We now present a group of theorems which have a bearing on this problem and which are of independent interest. Since there are common themes in the proofs we begin by presenting some general facts.

**16.4 Holomorphic maps: some observations.** Suppose that $G$ is an open set and that $f \in \mathrm{H}(G)$. Let $a \in G$.

(1) Assume that $G$ is a region and that $f$ is non-constant. Then, in some $\mathrm{D}'(a;r)$, the function $f-f(a)$ is never zero (by the Identity theorem (15.8)).

(2) Let $f$ be one-to-one. Then $f'$ certainly cannot be identically zero and hence can only have isolated zeros (by 15.8, applied to $f'$).

(3) Choose $r$ such that $\overline{\mathrm{D}}(a;r) \subseteq G$ and suppose that $f - f(a)$ is non-zero on $\gamma^*$, where $\gamma = \gamma(a;r)$. Let $m := \inf\{\, |f(z) - f(a)| : z \in \gamma^* \,\}$. Then
   (i) $m > 0$ (by 3.24);
   (ii) for each $w \in \mathrm{D}(f(a);m)$, the functions $f - f(a)$ and $f - w$ have the same number of zeros, counted according to multiplicity (by Rouché's theorem (15.14)): for $z \in \gamma^*$,

$$|f(z) - f(a)| \geqslant m > |f(a) - w| = |(f(a) - f(z)) + (f(z) - w)|.$$

**16.5 Theorem.** Suppose that $f$ is holomorphic and one-to-one in an open set $G$. Then $f$ is conformal in $G$.

**Proof** Assume, for a contradiction, that there exists $a \in G$ such that $f'(a) = 0$. Choose $r$ such that $\overline{\mathrm{D}}(a;r) \subseteq G$ and such that $f'$ is never zero in $\mathrm{D}'(a;r)$. This is possible by 16.4(2). Let $w \in \mathrm{D}'(f(a);m)$, where $m$ is as in 16.4(3). By 16.4(3)(ii), $f - f(a)$ and $f - w$ have the same number of zeros in $\mathrm{D}(a;r)$. The function $f - f(a)$ has a zero of order at least two at $a$ (by 15.3). On the other hand, $f - w$ cannot have two distinct zeros, since $f$ is one-to-one, and cannot have a zero of order greater than one, since $f - w$ and $(f - w)'$ cannot both be zero at any point in $\mathrm{D}(a;r)$. □

**16.6 Open mapping theorem.** Suppose that $f$ is holomorphic and non-constant in an open set $G$. Then $f(G)$ is open.

**Proof** Fix $a \in G$. Choose $r$ and $m$ as in 16.4(3). Observe that $f - w$ has a zero at a point $b$ if and only if $f(b) = w$, and that this implies that $w$ lies in the image of $f$. By 16.4(1) and 16.4(3)(ii), $f - w$ has at least one zero in $\mathrm{D}(a;r)$ whenever $w \in \mathrm{D}(f(a);m)$. Hence $f(a) \in \mathrm{D}(f(a);m) \subseteq f(\mathrm{D}(a;r)) \subseteq f(G)$. The result now follows from the definition of an open set. □

A special case of the following result was given in 7.16, where we considered holomorphy of a branch of the logarithm.

**16.7 Inverse function theorem.** Let $G$ be an open set and let $f$ be holomorphic and one-to-one in $G$. Then $f^{-1}$ is holomorphic in $f(G)$.

**Proof** Since $f$ is one-to-one, there exists an inverse map $g := f^{-1} \colon f(G) \to G$. Let $b = f(a) \in f(G)$. Then $a = g(b)$. We claim that $g$ is continuous. We prove this by applying the Open mapping theorem to $f$ in $\mathrm{D}(a;\varepsilon) \subseteq G$ to obtain $\delta > 0$ such that $\mathrm{D}(b;\delta) \subseteq f(\mathrm{D}(a;\varepsilon)) = g^{-1}(\mathrm{D}(a;\varepsilon))$. This is the statement of the $\varepsilon$-$\delta$ definition of continuity of $g$ at $b$ expressed in shorthand.

By Theorem 16.5, $f'(g(b)) \neq 0$. Then

$$\frac{g(w)-g(b)}{w-b} = \frac{g(w)-g(b)}{f(g(w))-f(g(b))} \to \frac{1}{f'(g(b))}$$

as $g(w) \to g(b)$. But $w \to b$ forces $g(w) \to g(b)$, since $g$ is continuous. □

**16.8 Conformality of invertible maps.** Let $G$ be a region and let $f \in \mathrm{H}(G)$. Suppose that $f$ maps $G$ one-to-one onto the region $\widehat{G} := f(G)$, so there exists a well-defined inverse map $f^{-1} \colon \widehat{G} \to G$. Then we have

- $f$ is conformal (by 16.5);
- $f^{-1}$ is conformal (by 16.5 and the Inverse function theorem).

There is a partial converse: if $f$ is conformal in a region $G$, then $f$ is locally one-to-one in $G$ (Exercise 16.5).

The next result greatly improves on one obtained in Exercise 2.13.

**16.9 Example (conformal mappings of the unit disc).** Suppose that $f$ is a one-to-one map of $\mathrm{D}(0;1)$ onto$\mathrm{D}(0;1)$ which is conformal in $\mathrm{D}(0;1)$ and suppose that $f(\alpha) = 0$ $(\alpha \in \mathrm{D}(0;1))$. We claim that, for some real constant $\lambda$,

$$f(z) = \mathrm{e}^{\mathrm{i}\lambda}\phi_\alpha(z) \quad (z \in \mathrm{D}(0;1)),$$

where $\phi_\alpha(z) := (z-\alpha)/(\overline{\alpha}z - 1)$. Recall from Exercise 2.13 that the Möbius transformation $\phi_\alpha$ is its own inverse and maps $\mathrm{D}(0;1)$ onto itself. To prove the claim about $f$ we consider the composite function $h = f \circ \phi_\alpha$. Then $h$ maps $\mathrm{D}(0;1)$ one-to-one onto itself, since both $f$ and $\phi_\alpha$ do. By the Inverse function theorem, $h^{-1}$ is also a holomorphic map of $\mathrm{D}(0;1)$ to itself. Schwarz' lemma (16.3) applies to each of $h$ and $h^{-1}$ to give

$$|h(w)| \leqslant |w| \quad \text{and} \quad |w| \leqslant |h(w)| \qquad (|w| < 1).$$

Together, these inequalities show that equality holds, so that, by the final statement in Schwarz' lemma, $h$ is a constant of modulus one. That is, $f(\phi_\alpha(w)) = e^{i\lambda}w$ for $w \in D(0;1)$, where $\lambda$ is a real constant. Recalling that $\phi_\alpha = \phi_\alpha^{-1}$, we deduce the required form for $f$.

**16.10 Do conformal mappings exist? The Riemann mapping theorem.** It is by no means obvious that it is even theoretically possible to construct a conformal mapping from a region with a complicated, spiky boundary onto a civilized region such as $D(0;1)$ or vice versa. The definitive theorem about this, the **Riemann mapping theorem**, is very striking: Let $G$ be a simply connected region with $G \neq \mathbb{C}$. Then there exists a one-to-one conformal mapping $f$ from $G$ onto $D(0;1)$ with $f^{-1}\colon D(0;1) \to G$ also conformal.

It is worth noting that in each of our worked examples in Chapter 8, the function we defined not only took one prescribed region, $G$, onto another, $\widehat{G}$, but also mapped the boundary of $G$ onto the boundary of $\widehat{G}$ (in $\widetilde{\mathbb{C}}$ in certain cases). Such an extension to the boundary is important in many applications, as our discussion in Chapter 23 indicates. In general, whether or not a conformal mapping on a region $G$ extends to a continuous function on $G \cup \partial G$ depends on the topological nature of the boundary.

## Exercises

16.1 Use Exercise 14.5 to obtain an alternative proof of the Local maximum modulus theorem.

16.2 Let $f$ be holomorphic in $\mathbb{C}$. Prove, by considering a suitable exponential, that the condition that $f(z)$ be real when $|z| = 1$ forces $f$ to be constant.

16.3 Let $G$ be the square region $\{\, z \in \mathbb{C} : |\mathrm{Re}\, z| < 1, |\mathrm{Im}\, z| < 1 \,\}$. Suppose that $f$ is continuous on $\overline{G}$ and holomorphic in $G$ and such that $f(z) = 0$ when $\mathrm{Re}\, z = 1$. By considering $g$ defined by

$$g(z) = f(z)f(iz)f(-z)f(-iz),$$

prove that $f$ is identically zero in $\overline{G}$.

16.4 Let $f \in H(D(0;R))$ smd let $M(r) := \sup\{\, |f(z)| : |z| = r \,\}$ $(r < R)$. Prove that $M(r) \leqslant M(s)$ for $r < s < R$, with strict inequality if $f$

is non-constant. Prove also that if $f$ is a polynomial of degree $n$ then $M(r)r^{-n} \geqslant M(s)s^{-n}$ when $0 < r < s < R$.

16.5 Use the facts given in 16.4 to prove that, if $f$ is conformal in a region $G$, then for each $a \in G$ there exists $r > 0$ such that the restriction of $f$ to $\mathrm{D}(a;r)$ is one-to-one.

16.6 Let $f$ satisfy the conditions for Schwarz' lemma (16.3). Complete the proof given there by considering the case of equality. Prove further that $|f'(0)| \leqslant M/R$.

16.7 Consider $\Pi^+ = \{\, z \in \mathbb{C} : \operatorname{Im} z > 0 \,\}$ and let $a \in \Pi^+$. Suppose that $F\colon \Pi^+ \to \Pi^+$ is holomorphic. Prove that, for all $z \in \Pi^+$,

$$\left|\frac{F(z) - F(a)}{F(z) - \overline{F(a)}}\right| \leqslant \left|\frac{z-a}{z-\bar{a}}\right| \quad \text{and} \quad |F'(a)| \leqslant \frac{\operatorname{Im} F(a)}{\operatorname{Im} a}.$$

(Hint: consider the composite function $f = \phi \circ F \circ \phi^{-1}$, where $\phi(z) = (z-a)/(z-\bar{a})$ and apply Schwarz' lemma.)

# 17 Singularities

The time has come to face the fact that many commonplace functions fail to be holomorphic at isolated points, or worse. Laurent's theorem (17.3) provides a very satisfactory substitute for Taylor's theorem for functions holomorphic except at isolated singular points.

## Laurent's theorem

**17.1 Binomial expansions again.** For $|z| < 1$, we can expand $(1-z)^{-1}$ in positive powers of $z$:

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n \quad (|z| < 1)$$

(recall 6.2). For $|z| > 1$ the series on the right-hand side no longer converges. Instead, we note that $|z| > 1$ if and only if $|1/z| < 1$ so that

$$\frac{1}{1-z} = -\frac{1}{z}\,\frac{1}{(1-1/z)} = -\sum_{n=0}^{\infty} z^{-n-1} = -\sum_{m=-\infty}^{-1} z^m \quad (|z| > 1).$$

In the same way, we can expand $(a-z)^{-1}$ as a series in positive powers of $z$ if $|z| < |a|$ and as a series in negative powers if $|z| > |a|$.

As an example, note that, in the annulus $\{\, z : 1 < |z| < 3 \,\}$, we have

$$\frac{4}{(1-z)(z+3)} = \frac{1}{1-z} + \frac{1}{z+3} = \sum_{n=-\infty}^{-1} z^n + \sum_{n=0}^{\infty} (-1)^n 3^{-n-1} z^n.$$

**17.2 Double-ended series.** By definition, a series $\sum_{n=-\infty}^{\infty} a_n$ converges (to $s = s_1 + s_2$) if $\sum_{n=0}^{\infty} a_n$ converges (to $s_1$) and $\sum_{n=1}^{\infty} a_{-n}$ converges (to $s_2$).

If $f$ is holomorphic in a disc $\mathrm{D}(a;r)$ except at $a$ itself, where something nasty happens, then we cannot hope for a power series expansion

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n$$

since the power series on the right-hand side behaves decently at $a$, by Theorem 6.8. Motivated by the examples in 17.1, we aim for a double-ended series: for $f \in \mathrm{H}(\mathrm{D}'(a;r))$ and seek to show that $f(z)$ can be expanded as a **Laurent series**

$$\sum_{n=-\infty}^{\infty} c_n(z-a)^n$$

valid for $0 < |z-a| < r$. We shall prove, slightly more generally, that a function holomorphic in an annulus has an expansion of this type.

**17.3 Laurent's theorem (series expansion in an annulus).** Let

$$A = \{z \in \mathbb{C} : R < |z-a| < S\} \quad (0 \leqslant R < S \leqslant \infty)$$

and let $f \in \mathrm{H}(A)$. Then

$$f(z) = \sum_{n=-\infty}^{\infty} c_n(z-a)^n \quad (z \in A),$$

where

$$c_n = \frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{f(w)}{(w-a)^{n+1}}\, \mathrm{d}w,$$

with $\gamma = \gamma(a;r)$ $(R < r < S)$ [or any closed path in $A$ homotopic to $\gamma(a;r)$].

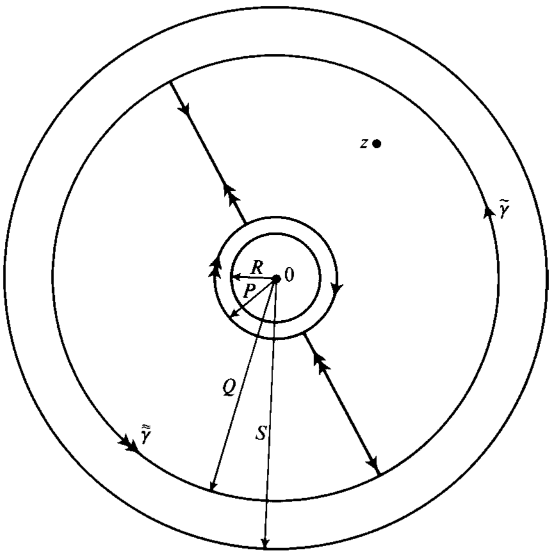



**Figure 17.1** The proof of Laurent's theorem

**Proof** By changing the origin if necessary, we may assume that $a = 0$. Fix $z \in A$ and choose $P$ and $Q$ such that $R < P < |z| < Q < S$. Let $\tilde{\gamma}$ and $\tilde{\tilde{\gamma}}$ be as shown in Fig. 17.1. Then

$$f(z) = \frac{1}{2\pi \mathrm{i}} \int_{\tilde{\gamma}} \frac{f(w)}{w-z}\,\mathrm{d}w \quad \text{(by 13.1)}$$

and

$$0 = \frac{1}{2\pi \mathrm{i}} \int_{\tilde{\tilde{\gamma}}} \frac{f(w)}{w-z}\,\mathrm{d}w \quad \text{(by Cauchy's theorem I (11.6))}.$$

Adding, and noting that the integrals along the line segments in $\tilde{\tilde{\gamma}}$ cancel, we have

$$\begin{aligned} f(z) &= \frac{1}{2\pi \mathrm{i}} \int_{\gamma(0;Q)} \frac{f(w)}{w-z}\,\mathrm{d}w - \frac{1}{2\pi \mathrm{i}} \int_{\gamma(0;P)} \frac{f(w)}{w-z}\,\mathrm{d}w \\ &= \frac{1}{2\pi \mathrm{i}} \int_{\gamma(0;Q)} \sum_{n=0}^{\infty} \frac{z^n}{w^{n+1}} f(w)\,\mathrm{d}w - \frac{1}{2\pi \mathrm{i}} \int_{\gamma(0;P)} \sum_{m=0}^{\infty} -\frac{w^m}{z^{m+1}} f(w)\,\mathrm{d}w, \end{aligned}$$

using the appropriate binomial expansions (we have $|z/w| < 1$ for $w \in \gamma(0;Q)^*$ and $|w/z| < 1$ for $w \in \gamma(0;P)^*$). We now invoke Theorem 14.2 [uniform convergence] to justify interchange of summation and integration (cf. the proof of Taylor's theorem (14.4)). This gives

$$f(z) = \sum_{n=0}^{\infty} \left( \frac{1}{2\pi \mathrm{i}} \int_{\gamma(0;Q)} \frac{f(w)}{w^{n+1}}\,\mathrm{d}w \right) z^n + \sum_{m=0}^{\infty} \left( \frac{1}{2\pi \mathrm{i}} \int_{\gamma(0;P)} f(w) w^m\,\mathrm{d}w \right) z^{-m-1}.$$

We now put $n = -m-1$ in the second sum. Finally, we use the Deformation theorem to replace $\gamma(0;Q)$ and $\gamma(0;P)$ by $\gamma$ as in the statement of Laurent's theorem. □

**17.4 Uniqueness of the Laurent expansion.** Let $f \in \mathrm{H}(A)$, where $A$ is the annulus $\{ z \in \mathbb{C} : R < |z-a| < S \}$ $(0 \leqslant R < S \leqslant \infty)$ and suppose that

$$f(z) = \sum_{n=-\infty}^{\infty} d_n (z-a)^n \quad (z \in A).$$

Then $d_n = c_n$ for all $n \in \mathbb{Z}$, where $c_n$ is as in 17.3. For the proof, we may assume $a = 0$. Choose $r$ such that $R < r < S$. Then

$$\begin{aligned} 2\pi \mathrm{i} c_n &= \int_{\gamma(0;r)} f(w) w^{-n-1}\,\mathrm{d}w \\ &= \int_{\gamma(0;r)} \sum_{k=-\infty}^{\infty} d_k w^{k-n-1}\,\mathrm{d}w \\ &= \int_{\gamma(0;r)} \sum_{k=0}^{\infty} d_k w^{k-n-1}\,\mathrm{d}w + \int_{\gamma(0;r)} \sum_{m=1}^{\infty} d_{-m} w^{-m-n-1}\,\mathrm{d}w \end{aligned}$$

Summation and integration can be interchanged to give

$$2\pi \mathrm{i} c_n = \sum_{k=-\infty}^{\infty} d_k \int_{\gamma(0;r)} w^{k-n-1}\,\mathrm{d}w = 2\pi \mathrm{i} d_n,$$

by the Fundamental integral (10.4). For the justification, consider the two integrals separately and appeal to the Interchange theorem (14.2 [or, via uniform convergence, 14.19]).

Uniqueness allows us relate Taylor series and Laurent series when the former exist. Suppose that $f$ is holomorphic in $\mathrm{D}(a;S)$. It has a Taylor expansion there and also has a Laurent expansion in $\mathrm{D}'(a;S)$. The uniqueness of the Laurent coefficients forces these expansions to coincide in $\mathrm{D}'(a;S)$ (with $c_n = 0$ for all $n < 0$).

**17.5 Computation of Laurent expansions.** We have already advocated using known expansions wherever possible to find Taylor expansions. This strategy relies on the uniqueness of the coefficients. As a fall-back, we may compute the higher-order derivatives $f^{(n)}(a)$ $(n \geqslant 0)$.

For a Laurent expansion, the method of known expansions is even more important. Cauchy's formula for the $n$th derivative is certainly not available for negative $n$ and calculating the Laurent coefficients directly from the formula

$$c_n = \frac{1}{2\pi \mathrm{i}} \int_{\gamma(a;r)} \frac{f(w)}{(w-a)^{n+1}}\,\mathrm{d}w$$

would be tedious, or impossible. So in practice we combine simple known expansions (Taylor or Laurent) to obtain Laurent expansions of more complicated functions. In applications, often only the first few terms in a Laurent expansion are needed. We shall see in the next chapter that the coefficient $c_{-1}$ in a Laurent expansion is of special significance. The reason for this can be traced back to the exceptional case, $n = -1$, in the Fundamental integral.

## 17.6 Examples (Laurent expansions).

### Tactical tip

⊙ Remember the examples in 6.3 and also that factors such as $(1-w)^{-2}$ can be expanded binomially for $|w| < 1$ (derivation by differentiation of the geometric series for $(1-z)^{-1}$)); recall 6.9. The same result can be arrived at by multiplication of series, but this is more laborious.

- **Exploiting binomial expansions** $f(z) = 1/(z(1-z)^2)$ is holomorphic in $A_1 = \{ z : 0 < |z-1| < 1 \}$ and in $A_2 = \{ z : |z-1| > 1 \}$. For $z \in A_1$, the binomial expansion gives

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

So, for $0 < |z-1| < 1$, the Laurent expansion for $f(z)$ is

$$f(z) = \sum_{n=-2}^{\infty} (-1)^n (z-1)^n.$$

For $|z-1| > 1$ we write $f(z)$ as $(z-1)^{-3}(1+1/(z-1))^{-1}$. This expands binomially to give

$$f(z) = \sum_{n=-\infty}^{-3} (-1)^{n+1}(z-1)^n \quad (|z| > 1).$$

- **Inverting a known expansion** The function $\operatorname{cosec} z$ is holomorphic except at $z = k\pi$ ($k \in \mathbb{Z}$) and so has a Laurent expansion $\sum_{n=-\infty}^{\infty} c_n z^n$ valid for $0 < |z| < \pi$. We have

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots = z\Big(1 - \frac{z^2}{3!} + h(z)\Big),$$

where all the terms after the first two have been amalgamated to form the function $h(z)$. This function $h$ is defined by a convergent power series, and so is holomorphic. Near 0, the dominant term is that in $z^4$ and we have $|h(z)| \leqslant K\left|z^4\right|$ for some constant $K$; we shall use the conventional **O**-notation for this and write $h(z) = \mathbf{O}(z^4)$. Then

$$\operatorname{cosec} z = \frac{1}{z}\Big(1 - (\frac{z^2}{3!} + \mathbf{O}(z^4))\Big)^{-1} = \frac{1}{z}\Big(1 + \frac{z^2}{3!} + \mathbf{O}(z^4)\Big) \qquad \text{for small } |z|.$$

Here we have used the standard expansion $(1-w)^{-1} = 1+w+w^2+\dots$, valid for $|w| < 1$, with $w = z^2/3! + h(z)$. Then 17.4 implies that $c_n = 0$ for all

$n < -1$, $c_{-1} = 1$, $c_1 = 1/6$. By taking more terms in the above expansions, we could compute $c_3, c_5, \ldots$. Trivially, $c_{2k} = 0$ for all integers $k$.

- **Changing the base point** To find the Laurent expansion of $\operatorname{cosec} z$ about $z = k\pi$ $(k \in \mathbb{Z},\ k \neq 0)$, we want to expand in powers of $w = z - k\pi$. By the addition formula for the sine function,

$$\operatorname{cosec} z = (-1)^k \operatorname{cosec}(z - k\pi) = (-1)^k \sum_{n=-\infty}^{\infty} c_n (z - k\pi)^n,$$

where the coefficients $c_n$ are as in the expansion about 0.

- **Beware bogus expansions!** Replacing $z$ by $1/z$ in the expansion for $\operatorname{cosec} z$ above appears to give

$$\operatorname{cosec}(1/z) = z\left(1 + \frac{1}{3!z^2} + \mathbf{O}(z^{-4})\right).$$

This is *not* valid. The substitution of $1/z$ for $z$ overlooks the restriction to small $|z|$ imposed above to validate inverting the sine expansion. In fact there is no Laurent expansion about 0, since there is no punctured disc $\mathrm{D}'(0;\varepsilon)$ in which $\operatorname{cosec}(1/z)$ is holomorphic. (In the terminology of 17.8 below, the singularity at 0 is not isolated.)

- **Multiplying known expansions** $\cot z = \dfrac{\cos z}{\sin z}$ is holomorphic for $0 < |z| < \pi$. Near 0,

$$\begin{aligned}\cot z &= \left(1 - \frac{z^2}{2!} + \mathbf{O}(z^4)\right)\left(\frac{1}{z} + \frac{z}{6} + \mathbf{O}(z^3)\right)\\ &= \frac{1}{z}\left(1 + z^2\left(-\frac{1}{2} + \frac{1}{6}\right) + \mathbf{O}(z^4)\right),\end{aligned}$$

using the expansion above; remember that multiplication of convergent power series is permissible, by 14.10. Hence

$$\cot z = \frac{1}{z} - \frac{z}{3} + \mathbf{O}(z^3) \quad (0 < |z| < \pi).$$

Our second example of this type is drawn from fluid mechanics. Let

$$f(z) = \left(z + \frac{c^2}{z}\right)\left(1 - \frac{a^2}{z^2}\right)^2\left(1 - \frac{c^2}{z^2}\right)^{-1} \quad (|z| > c),$$

where $a$ and $c$ are positive constants. Expanding $\left(1 - (c^2/z^2)\right)^{-2}$ binomially for large $|z|$,

$$f(z) = \left(z + \frac{c^2}{z}\right)\left(1 - \frac{a^2}{z^2}\right)^2\left(1 + \frac{c^2}{z^2} + \frac{c^4}{z^4} + \ldots\right).$$

This is a case where there are infinitely many negative powers. The coefficient of any given $z^n$ $(n < 0)$ could be computed by collecting together terms. For example, the coefficient of $z^{-1}$ is $c^2 - 2a^2 + c^2$.

**17.7 Estimating Laurent coefficients.** Suppose that $f$ is holomorphic in an annulus $\{ z : R < |z| < S \}$. Let $f$ have Laurent expansion $f(z) = \sum_{n=-\infty}^{\infty} c_n z^n$, where $2\pi i c_n = \int_{\gamma(0;r)} f(w) w^{-n-1}\, dw$ $(R < r < S)$. Estimating in the same way as in 14.7, we have

$$|c_n| \leqslant r^{-n} \sup\{ |f(z)| : |z| = r \}.$$

Two special cases are worth noting.

- Assume that $f$ is holomorphic for $|z| > R$ and that $f$ is bounded ($|f(z)| \leqslant M$ say). Then $|c_n| \leqslant M r^{-n}$ for all $r > R$. This forces $c_n = 0$ for all $n > 0$, and so the Laurent expansion of $f$ takes the form $\sum_{n=-\infty}^{0} c_n z^n$.
- Assume that $f$ is defined and holomorphic in a punctured disc $D'(0; S)$. This time we can take $r$ arbitrarily small so that our estimate gives $c_n = 0$ for $n < 0$. Therefore $f(z) = \sum_{n=0}^{\infty} c_n z^n$ for $0 < |z| < S$. In addition, if we *define* $f(0) = c_0$, we obtain

$$f(z) = \sum_{n=0}^{\infty} c_n z^n \quad (|z| < S),$$

and this power series is holomorphic in $D(0; S)$. We pursue these ideas in 17.15, where we consider removable singularities.

## Singularities

**17.8 Definitions (singularities).** We say $a$ is a **regular point** (of $f$) if $f$ is holomorphic at $a$ (that is, $f \in H(D(a; r))$ for some $r > 0$; see 5.7). A point $a$ is a **singularity** of $f$ if $a$ is a limit point of regular points which is not itself regular.

If $a$ is a singularity of $f$ and $f$ is holomorphic in some punctured disc $D'(a; r)$ $(r > 0)$, then $a$ is an **isolated singularity**; otherwise $a$ is a **non-isolated (essential) singularity**.

**17.9 Classification of isolated singularities.** Suppose that $f$ has an isolated singularity at $a$. Then $f$ is holomorphic in some annulus $\{ z : 0 < |z-a| < r \}$ and has there a unique Laurent expansion

$$f(z) = \sum_{n=-\infty}^{\infty} c_n(z-a)^n.$$

We may write

$$f(z) = \sum_{n=-\infty}^{-1} c_n(z-a)^n + \sum_{n=0}^{\infty} c_n(z-a)^n.$$

The second term on the right-hand side is holomorphic in $\mathrm{D}(a;r)$ and is in no way responsible for the singularity. This is caused by the first sum,

$$\sum_{n=-\infty}^{-1} c_n(z-a)^n,$$

which is known as the **principal part** of the Laurent expansion.

We classify isolated singularities according to the behaviour of the coefficients $c_n$ for $n < 0$. We stress that the feasibility of such a classification relies on both the *existence* and the *uniqueness* of the Laurent expansion. The point $a$ is said to be

- a **removable singularity** if $c_n = 0$ for all $n < 0$;
- a **pole of order** $m$ ($m \geqslant 1$) if $c_{-m} \neq 0$ and $c_n = 0$ for all $n < -m$;
- an **isolated essential singularity** if there does not exist $m$ such that $c_n = 0$ for all $n < -m$.

Poles of orders $1, 2, 3, \ldots$ are called **simple, double, triple,** $\ldots$.

**17.10 Examples (principal parts and singularities).** The examples that follow are preliminary ones to illustrate the definitions in situations where the Laurent expansion, or at least its principal part, can be written down easily.

- $(z-1)^{-2}$ is its own Laurent expansion about $z = 1$, where it has a double pole.
- We showed in 17.6 that

  $$\cot z = \frac{1}{z} - \frac{z}{3!} + \mathbf{O}(z^3) \quad (z \in \mathrm{D}'(0;\pi)).$$

  The principal part of the Laurent expansion about 0 is $z^{-1}$ and $\cot z$ has a simple pole at 0. Since $\cot(z - k\pi) = \cot z$ for each integer $k$, each singularity $k\pi$ of $\cot z$ is a simple pole.

- $(1-\cos z)z^{-2}$ is holomorphic except at $z=0$ where it is indeterminate. The Laurent expansion about 0 is

$$\frac{1}{2!}-\frac{z^2}{4!}+\frac{z^4}{6!}-\dots \quad (|z|>0).$$

The singularity at 0 is removable.

- For $0<|z|<\infty$,

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

Hence $\sin(1/z)$ has an isolated essential singularity at 0.

It should be clear that computing the Laurent coefficients is an arduous way of classifying the singularities of even relatively simple functions. Fortunately there is a much more efficient method. The clue to it lies in the observation that, if a holomorphic function has an isolated zero at the point $a$, then its reciprocal has an isolated singularity at $a$. The following technical result parallels that for zeros in 15.3. We leave the proof as an exercise; it is very like that given in 15.3.

**17.11 Characterization theorem for poles of order $m$.** Let $f\in \mathrm{H}(\mathrm{D}'(a;r))$. Then $f$ has a pole of order $m$ at $a$ if and only if

$$\lim_{z\to a}(z-a)^m f(z)=D, \quad \text{where } D \text{ is a finite non-zero constant.}$$

**17.12 Theorem (poles and zeros).** Suppose that $f$ is holomorphic in some open disc $\mathrm{D}(a;r)$. Then $f$ has a zero of order $m$ at $a$ if and only if $1/f$ has a pole of order $m$ at $a$.

**Proof** Suppose $1/f$ has a pole at $a$. This requires that $1/f$ be holomorphic in some punctured disc with centre $a$. Therefore the zero of $f$ at $a$ cannot be non-isolated, by the Identity theorem (15.7), and so 15.3 is applicable. Conversely, a zero $a$ of $f$ of order $m$ is necessarily isolated, so $1/f$ is holomorphic in some punctured disc $\mathrm{D}'(a;r)$. The result now follows from 15.3 and 17.11 and the algebra of limits. □

We leave as an exercise the proof of the following very useful consequence of the theorems characterizing zeros and poles.

**17.13 Cancellation and coalescence of zeros and poles.** Suppose that $f$ has a pole of order $m$ at $a$.

(1) Suppose that $g \in \mathrm{H}(\mathrm{D}(a;r))$ for some $r > 0$. Then at $a$ the function $fg$ has
- a pole of order $m$ if $g(a) \neq 0$;
- a pole of order $m-n$ if $g$ has a zero of order $n < m$ at $a$;
- a removable singularity if $g$ has a zero of order $n \geqslant m$ at $a$.

(2) Suppose that $g$ has a pole of order $n$ at $a$. Then $fg$ has a pole of order $m+n$ at $a$.

**17.14 Examples (orders of poles).**

- $z \sin z$ has isolated zeros at $z = k\pi$ ($k \in \mathbb{Z}$), all being simple except the zero at $0$, which is double. Therefore $1/(z \sin z)$ has a simple pole at $k\pi$ ($0 \neq k \in \mathbb{Z}$) and a double pole at $0$.
- Consider again $\cot z = \cos z / \sin z$. At the points $z = k\pi$ ($k \in \mathbb{Z}$), $\sin z$ has simple zeros and $\cos z \neq 0$. Therefore $\cot z$ has a simple pole at each point $k\pi$.

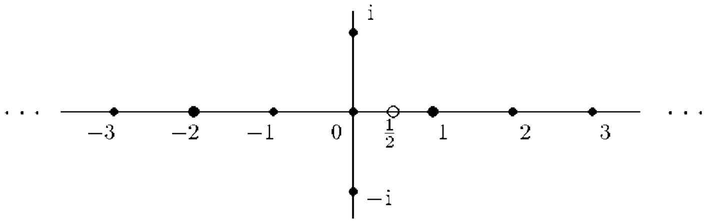



**Figure 17.2** Singularities of $f(z) = \dfrac{(z-1)\cos \pi z}{(z+2)(2z-1)(z^2+1)^3 \sin^2 \pi z}$

- Consider
$$f(z) = \frac{(z-1)\cos \pi z}{(z+2)(2z-1)(z^2+1)^3 \sin^2 \pi z}.$$
The denominator has a simple zero at $1/2$, zeros of order 3 at $\pm\mathrm{i}$, a double zero at each integer $k \neq -2$, and a triple zero at $-2$. The numerator has a simple zeros at 1 and at $(2k+1)/2$ for each integer $k$. Appealing to 17.13, we see that $f$ has triple poles at $-2$ and $\pm\mathrm{i}$, a double pole at $k$ ($k \in \mathbb{Z} \setminus \{1, -2\}$), a simple pole at 1, and a removable singularity at $1/2$. In Fig. 17.2, we have depicted by large circles the singularities arising from coalescence or cancellation of zeros in different factors in the expression $f(z)$.

**17.15 Removable singularities.** Let $f \in \mathrm{H}(\mathrm{D}'(a;r))$ and assume that $f$ has a removable singularity at $a$. We have a Laurent expansion $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ in $\mathrm{D}'(a;r)$. Then $f(z) \to c_0$ as $z \to a$ (by continuity of the right-hand side at $a$). By defining (or re-defining) $f(a)$ to be $c_0$ we arrive at

$$f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n \quad (|z-a| < r)$$

and so have made $f$ holomorphic in $\mathrm{D}(a;r)$, by 6.8. Thus a removable singularity is a non-event: $a$ ceases to be classified as a singularity once $f$ is correctly defined at $a$. The standard theorems about holomorphic functions—Cauchy's theorem, Liouville's theorem, etc.—can then be applied.

Suppose now that we have any function $f$ with an isolated singularity at $a$, so that there is a Laurent expansion

$$f(z) = \sum_{n=-\infty}^{-1} c_n(z-a)^n + \sum_{n=0}^{\infty} c_n(z-a)^n \quad (z \in \mathrm{D}'(a;r)).$$

By subtracting from $f(z)$ its principal part, namely $\sum_{n=-\infty}^{-1} c_n(z-a)^n$, we convert the singularity at $a$ to a removable one, which can then be removed.

**17.16 Behaviour near a non-removable isolated singularity.** Let $f$ have an isolated singularity at $a$ and have Laurent expansion

$$f(z) = \sum_{n=-\infty}^{\infty} c_n(z-a)^n \quad (0 < |z-a| < r).$$

- **Pole** Suppose $f$ has a pole at $a$. It is immediate from 17.11 that $|f(z)| \to \infty$ as $z \to a$.
- **Essential singularity** Suppose that $f$ has an isolated essential singularity at $a$. Let $w$ be any complex number. Then there exists a sequence $\{a_n\}$ such that $a_n \to a$ and $f(a_n) \to w$. This is the **Casorati–Weierstrass theorem.** For an outline of the proof, see Exercise 17.15.

  A more spectactular and much deeper result, due to Picard, asserts that in any $\mathrm{D}'(a;\varepsilon)$, $f$ actually assumes every complex value except, possibly, one. In the case of $\mathrm{e}^{1/z}$, which has an isolated essential singularity at 0, the exceptional value is 0.

**17.17 Non-isolated singularities.** We emphasize that in order that a function $f$ have a Laurent expansion about a point $a$ it is necessary that $f$ be holomorphic in some punctured disc $\mathrm{D}'(a;r)$ $(r > 0)$. This fails to happen whenever $a$ is a limit point of singularities of $f$.

- $\operatorname{cosec}(1/z)$ has singularities at $z = 0$ (where it is undefined) and at $1/(k\pi)$ $(0 \neq k \in \mathbb{Z})$ (where $\sin(1/z) = 0$). At $1/(k\pi)$ there is a simple pole (see 17.6). There is *no* punctured disc $\mathrm{D}'(0;r)$ in which $\operatorname{cosec}(1/z)$ is holomorphic since every such disc contains points of the form $1/(k\pi)$. Hence 0 is a non-isolated singularity.

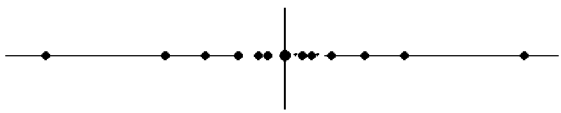


**Figure 17.3** Singularities of $\operatorname{cosec}(1/z)$

- $f(z) = z^{-3}(1+e^{1/z})^{-1}$ has singularities at 0 and at $1/((2k+1)\pi i)$ $(k \in \mathbb{Z})$, where $1+e^{1/z}$ has simple zeros. The point 0 is a limit point of singularities and so a non-isolated singularity. The fact that the factor $z^{-3}$ has a triple pole at 0 is irrelevant: $f$ itself has no Laurent expansion about 0.

## Meromorphic functions

We conclude this chapter by extending our investigation of singularities to the extended complex plane $\widetilde{\mathbb{C}}$.

**17.18 Singularities at $\infty$.** In 5.10 we briefly considered functions on $\widetilde{\mathbb{C}}$ and indicated that the inversion map $z \mapsto 1/z$ can be used to analyse what happens at or near $\infty$. Consider a function $f$ defined on some set $\{ z \in \mathbb{C} : |z| > r \}$ but not necessarily at the point $\infty$. As before, define $\widetilde{f}$ by

$$\widetilde{f}(w) = f(1/w) \quad (w \in \mathrm{D}'(0;1/r))$$

and let $\widetilde{f}(0) = f(\infty)$ if $f(\infty)$ is defined. We then transfer notions relating to $\widetilde{f}$ at 0 to obtain corresponding notions for $f$ at $\infty$. This device allows us to consider singularities, poles, ... at $\infty$. Here are some examples.

- For $f(z) = z^3$ we have $\widetilde{f}(w) = w^{-3}$, which has a triple pole at 0. Hence $f$ has a triple pole at $\infty$.

- $z^{-2}\sin z$ has a removable singularity at $\infty$.
- $\tan z$ has a non-isolated singularity at $\infty$. This can be seen directly, since $\infty$ is a limit point of the set $\{\,(2k+1)\pi/2 \,:\, k \in \mathbb{Z}\,\}$ of poles of $\tan z$.

**17.19 Definition (meromorphic function).** Let $G$ be an open subset of $\mathbb{C}$ or, more generally, of $\widetilde{\mathbb{C}}$. A complex-valued function which is holomorphic in $G$ except possibly for poles is said to be **meromorphic** in $G$.

**17.20 Theorem (meromorphic functions in $\widetilde{\mathbb{C}}$).**

(1) Let $f$ be holomorphic in $\widetilde{\mathbb{C}}$. Then $f$ is constant.

(2) Let $f$ be meromorphic in $\widetilde{\mathbb{C}}$. Then $f$ is a rational function.

**Proof** (1) The result follows immediately from Liouville's theorem (13.3) once we know that $f$ is bounded [which, as a complex-valued continuous function on the compact space $\widetilde{\mathbb{C}}$, it is]. An elementary proof of boundedness goes as follows. The function $\widetilde{f}$ is continuous on the compact subset $\overline{\mathbb{D}}(0;1)$ of $\mathbb{C}$, by 3.24. Consequently $f$ is bounded on $\{\,z \in \mathbb{C} \,:\, |z| \geqslant 1\,\} \cup \{\infty\}$. By 3.24, this time applied to $f$ itself, $f$ is bounded on $\{\,z \in \mathbb{C} \,:\, |z| \leqslant 1\,\}$.

(2) An infinite set of poles of $f$ would have a limit point in $\widetilde{\mathbb{C}}$ (see Exercise 3.15). Such a limit point would be a non-isolated singularity of $f$, and so could not be a pole. So $f$ has at most finitely many poles. Assume $f$ has a pole of order $m_k$ at $a_k \in \mathbb{C}$ ($k = 1, \ldots, N$) and a pole of order $m$ at $\infty$ (and let $m = 0$ if there is no pole at $\infty$). By the algebra of limits and 17.12,

$$\lim_{z \to a_n} \prod_{k=1}^{N} (z - a_k)^{m_k} z^{-m} f(z)$$

exists for each $n = 1, \ldots, N$. It follows that

$$g(z) = \prod_{k=1}^{N} (z - a_k)^{m_k} z^{-m} f(z)$$

has, at worst, removable singularities. Remove these (see 17.15) to make $g$ holomorphic in $\widetilde{\mathbb{C}}$. Applying (1) to $g$, we see that $g$ must be constant. □

Exercise 17.17 seeks an alternative proof of (2). In this, the singularities are cancelled out by subtraction of principal parts (see 17.15).

## Exercises

**Exercises from the text.** Fill in the details of the proof in 17.4 and prove the assertions in 17.13. Prove the Characterization theorem for poles (17.12).

17.1 Find the Laurent expansion of $(z^2-1)^{-2}$ valid for (i) $0<|z-1|<2$ and (ii) $|z+1|>2$.

17.2 (a) Let $f(z) = e^{z-1/z}$ $(0<|z|<\infty)$. Use the exponential series and 14.10 to write down an expression for the coefficient $c_n$ in the Laurent expansion for $f(z)$ valid for $0<|z|<\infty$.

(b) Prove that

$$\int_{\gamma(0;1)} \frac{e^{w-1/w}}{w^{n+1}}\, dw = i\int_0^{2\pi} \cos(n\theta - 2\sin\theta)\, d\theta.$$

(c) Deduce that

$$\frac{1}{2\pi}\int_0^{2\pi} \cos(n\theta - 2\sin\theta)\, d\theta = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!}.$$

17.3 Suppose that $f$ is continuous and bounded in $D(a;r)$ and that $f$ is holomorphic on $D'(a;r)$. Prove that $f \in H(D(a;r))$. (Hint: see 17.7.)

17.4 Suppose that, for $R<|z|<S$, $f(z) = g(z)+h(z)$, where $g$ is holomorphic for $|z|<S$ and $h$ is holomorphic and bounded for $|z|>R$. Let $\{c_n\}$ be the Laurent coefficients of $f$ in the annulus $R<|z|<S$. Prove, by considering the expansions of $g$ and $h$, that

$$g(z) = c + \sum_{n=1}^{\infty} c_n z^n \quad (|z|<S),$$

where $c$ is a constant. (Hint: use 17.7.)

17.5 Find the principal part of the Laurent expansion about 0 of

(i) $\dfrac{1}{z^2\sinh z}$, (ii) $\dfrac{1}{\sin^2 z}$, (iii) $(e^z-1)^{-2}$,

(iv) $e^{-1/z^4}$, (v) $\dfrac{e^z+1}{e^z-1}$, (vi) $(\sin z + \sinh z - 2z)^{-1}$.

In each case, specify the type of singularity the function has at 0.

17.6 Find the principal part of the Laurent expansion about the indicated point $a$ of

$$\text{(i) } \sec^2 z \quad (a = \pi/2), \qquad \text{(ii) } \frac{e^z - 1}{e^z + 1} \quad (a = \pi i),$$
$$\text{(iii) } (z^2 + b^2)^{-3} \quad (a = ib), \qquad \text{(iv) } \frac{e^{iz}}{(z^2 + b^2)^2} \quad (a = ib).$$

In each case, specify the type of singularity the function has at $a$.

17.7 The function $g$ is holomorphic in $\Pi^- = \{ z : \operatorname{Im} z < 0 \}$ and is such that $g(z + 2\pi) = g(z)$ for all $z \in \Pi^-$.

(i) Prove that there is a well-defined function $f$ defined on an unbounded annulus and such that $f(e^{iz}) = g(z)$.

(ii) Deduce that there exists an expansion

$$g(z) = \sum_{n=-\infty}^{\infty} c_n e^{inz} \quad (\operatorname{Im} z < 0).$$

17.8 (a) Locate and classify the singularities in $\mathbb{C}$ of the following functions:

$$\text{(i) } \frac{z}{(z^2 - 1)}, \qquad \text{(ii) } \frac{1}{z(3 - z)},$$
$$\text{(iii) } \frac{1}{z^3(z^2 + 1)}, \qquad \text{(iv) } \frac{1}{1 + z^4},$$
$$\text{(v) } \frac{1}{1 - z + z^2}, \qquad \text{(vi) } \frac{1}{(z^2 + z + 1)^3},$$
$$\text{(vii) } \frac{z}{(z + 1)^2(z^2 - 3z + 2)}, \qquad \text{(viii) } \left(\frac{z - 1}{z + 1}\right)^3.$$

(Hint: you will need to be familiar with the facts in 1.7.)

(b) Locate and classify the singularities in $\mathbb{C}$ of the following functions:

$$\text{(i) } \frac{1}{1 - e^z}, \qquad \text{(ii) } \frac{1}{(1 - \cos z)^2}, \qquad \text{(iii) } \tan^2 z, \qquad \text{(iv) } \frac{\operatorname{cosec} \pi z}{z^2 + 1}.$$

(Hint: you will need to be familiar with the facts in 7.10).)

17.9 Locate and classify the singularities in $\mathbb{C}$ of the following functions:

$$\text{(i) } \frac{z^2 + 1}{z^4 - 1}, \quad \text{(ii) } \frac{1}{z^4 \sin z}, \quad \text{(iii) } \frac{\cot \pi z}{z^6 - 1}, \quad \text{(iv) } \frac{z}{\sinh^2 z},$$
$$\text{(v) } \frac{z}{1 - e^z}, \quad \text{(vi) } \frac{z \sin z}{\cos z - 1}, \quad \text{(vii) } \frac{e^{iz}}{\cosh z}, \quad \text{(viii) } \frac{e^{1/z}}{z}.$$

(Hint: this is principally an exercise on the use of 17.13.)

17.10 Locate and classify the singularities in $\mathbb{C}$ of

$$\text{(i)}\ \frac{1}{(\pi+z)\sin z}, \qquad \text{(ii)}\ \frac{1}{(\pi+z)\sin z}-\frac{1}{\pi z}.$$

17.11 Locate and classify the singularities in $\mathbb{C}$ of

$$\frac{1}{\sin(\pi(z-1))\sin(\pi/(z+1))}.$$

17.12 Assume that $f$ has a pole of order $m$ at $a$ and that $g$ has a pole of order $n$ at $a$. What kind of singularity at $a$ is it possible for (i) $f+g$ and (ii) $f\circ g$ to have? Give examples to show that all the possibilities that you list can occur.

17.13 Let $g$ be holomorphic in $\mathbb{C}$ and assume there exists a finite constant $M$ such that

$$|g(z)| \leqslant M|\sin z| \quad (z \in \mathbb{C}).$$

Prove that $g(z) = K\sin z$ on $\mathbb{C}$, for some constant $K$ with $|K| \leqslant M$. (Hint: consider $g(z)/\sin z$.)

17.14 Let $G$ be a bounded region in $\mathbb{C}$ and let $S$ be a closed set contained in $G$.
(i) Show that there does not exist a function which has an infinite number of poles in $S$ and is otherwise holomorphic in $G$.
(ii) Show that there does not exist a function $f$ which is holomorphic in $G$ and is such that $1/f$ has an infinite number of poles in $S$.
(Hint: you will need the Identity theorem and the Bolzano–Weierstrass theorem.)

17.15 Suppose that $f$ is holomorphic in a punctured disc, centre $a$. Let $w \in \mathbb{C}$ be given. Suppose that there exist $\varepsilon > 0$ and $r > 0$ such that $|f(z)-w| \geqslant \varepsilon$ for all $z \in \mathrm{D}'(a;r)$. By considering the function $(f-w)^{-1}$, prove that $f$ cannot have an essential singularity at $a$. Deduce the Casorati–Weierstrass theorem stated in 17.16.

17.16 Determine the type of singularity that each of the following functions has (a) at 0 and (b) at $\infty$:

$$\text{(i)}\ \left(z+z^{-1}\right)^{-1}, \qquad \text{(ii)}\ z^2 e^{1/z}, \qquad \text{(iii)}\ \cot z - z^{-1},$$

$$\text{(iv)}\ (z-\sin z)^{-1}, \qquad \text{(v)}\ \operatorname{cosec}(\sin z), \qquad \text{(vi)}\ \frac{z^2}{\cosh z - \cos z}.$$

17.17 Give an alternative proof of Theorem 17.20(2) by considering the function obtained from $f$ by subtracting from it the sum of the principal parts of the Laurent expansion about each of its poles.

17.18 Let $f$ be holomorphic in $\mathbb{C}$.

(i) Prove that $f$ has a removable singularity at $\infty$ if and only if $f$ is constant.

(ii) Prove that $f$ has a pole of order $m$ at $\infty$ if and only if $f$ is a polynomial of degree $m$.

17.19 Construct functions $f_k$ $(k = 1, \ldots, 6)$ such that

(i) $f_1$ is holomorphic in $\mathbb{C}$ except for simple poles at $\pm 1$ and $\pm \mathrm{i}$;

(ii) $f_2$ is holomorphic in $\mathbb{C}$ except for removable singularities at $\pm 1$;

(iii) $f_3$ is holomorphic in $\mathbb{C}$ except for a pole of order 4 at $2k+1$ $(k \in \mathbb{Z})$;

(iv) $f_4$ is holomorphic in $\mathbb{C}$ except for non-isolated singularities at $\pm 1$ and a set of simple poles;

(v) $f_5$ is holomorphic in $\tilde{\mathbb{C}}$ except for a pole of order 6 at $\infty$;

(vi) $f_6$ is holomorphic in $\tilde{\mathbb{C}}$ except for isolated essential singularities at 0, 1, $\infty$.

17.20 (This exercise assumes familiarity with uniform convergence.) Define

$$f(z) = \frac{1}{z} - 2z \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2\pi^2 - z^2}.$$

(i) Prove that $f$ is holomorphic in $G := \mathbb{C} \setminus \{ n\pi : n \in \mathbb{Z} \}$ by showing that the series converges uniformly on any disc $\mathrm{D}(a;r) \subset G$ and applying the result in Exercise 14.7.

(ii) Prove that $f$ has a simple pole at each point $k\pi$ $(k \in \mathbb{Z})$. (Hint: split $f(z)$ into two parts: a finite sum having a pole at $k\pi$ and an infinite sum holomorphic at $k\pi$.)

(iii) Deduce that $f(z) = \operatorname{cosec} z$ for all $z \in \mathbb{C}$.

# 18 Cauchy's residue theorem

This chapter continues the study of singularities and extends in a significant way the techniques stemming from Cauchy's theorem for evaluating the integral of a function round a contour.

## Residues and Cauchy's residue theorem

The following lemma comes directly out of Laurent's theorem and the Deformation theorem: it gives an integral formula for the Laurent coefficient $c_{-1}$. The result is of sufficient importance for us to record it explicitly.

**18.1 Lemma (Integration round a pole).** Let $f$ be holomorphic inside and on a positively oriented contour $\gamma$ except at the point $a$ inside $\gamma$, where it has a pole of order $m$. Let

$$f(z) = \sum_{n=-m}^{\infty} c_n (z-a)^n$$

be the (unique) Laurent expansion of $f$ about $a$. Then

$$\int_{\gamma} f(z)\,\mathrm{d}z = 2\pi \mathrm{i} c_{-1}.$$

**18.2 Definition (residue).** Suppose that $f \in \mathrm{H}(\mathrm{D}'(a;r))$ and that $f$ has a pole at $a$. The **residue** of $f$ at $a$ is the (unique) coefficient $c_{-1}$ of $(z-a)^{-1}$ in the Laurent expansion of $f$ about $a$, and is denoted $\operatorname{res}\{f(z);a\}$.

**18.3 Cauchy's residue theorem.** Let $f$ be holomorphic inside and on a positively oriented contour except for a finite number of poles, $a_1, \ldots, a_N$ inside $\gamma$. Then

$$\int_{\gamma} f(z)\,\mathrm{d}z = 2\pi \mathrm{i} \sum_{k=1}^{N} \operatorname{res}\{f(z);a_k\}.$$

**Proof** Let $f_k(z)$ be the principal part of the Laurent expansion about $a_k$. Then

$$g := f - \sum_{k=1}^{N} f_k$$

has only removable singularities at $a_1, \ldots, a_N$; remove them (see 17.15). By Cauchy's theorem, $\int_\gamma g(z)\,\mathrm{d}z = 0$. Hence

$$\int_\gamma g(z)\,\mathrm{d}z = \int_\gamma f(z)\,\mathrm{d}z - \sum_{k=1}^{N} \int_\gamma f_k(z)\,\mathrm{d}z = 2\pi\mathrm{i} \sum_{k=1}^{N} \operatorname{res}\{f(z); a_k\},$$

by Lemma 18.1, applied to each $f_k$. □

**18.4 Example (function-finding).** Suppose that $f$ is holomorphic in $\mathbb{C}$ except for simple poles at the cube roots of unity, $1$, $\omega = \mathrm{e}^{2\pi\mathrm{i}/3}$, $\omega^2$, where it has residues $1$, $\alpha$ $(\neq 0)$, $\alpha^{-1}$, respectively. Suppose further that there exists a constant $K$ such that $|z^2 f(z)| \leqslant K$ for $|z| \geqslant 2$.

By Cauchy's residue theorem,

$$2\pi\mathrm{i}(1 + \alpha + \alpha^{-1}) = \int_{\gamma(0;R)} f(z)\,\mathrm{d}z \quad (R \geqslant 2).$$

By the Estimation theorem and the given growth bound,

$$|2\pi\mathrm{i}(1 + \alpha + \alpha^{-1})| \leqslant \int_0^{2\pi} |f(R\mathrm{e}^{\mathrm{i}\theta})|\, R\,\mathrm{d}\theta \leqslant 2\pi K/R,$$

Since $R$ can be arbitrarily large, this forces $1 + \alpha + \alpha^{-1} = 0$. Therefore there are two possibilities for $\alpha$, namely $\omega$ and $\omega^2$.

Subtracting off the principal parts of the Laurent expansions about the two poles, as in the proof of Cauchy's residue theorem, we obtain a function

$$g(z) := f(z) - \frac{1}{z-1} - \frac{\alpha}{z-\omega} - \frac{\alpha^{-1}}{z-\omega^2}$$

which has only removable singularities and so can be treated as holomorphic in $\mathbb{C}$. The growth bound on $f$ implies that $f(z)$ tends to zero as $|z| \to \infty$, and the same is true of $g(z)$. Combining this with 3.24, we see that $g$ is bounded in $\mathbb{C}$. By Liouville's theorem, $g$ is constant, and the constant must be zero. Put in the values for $\alpha$ found above and simplify. We deduce that

$$f(z) = 3(z^3 - 1)^{-1} \quad \text{or} \quad f(z) = 3z(z^3 - 1)^{-1}.$$

It is easily seen that each of these functions does meet the given conditions.

We could alternatively have cancelled out the singularities of $f$ by multiplying it by $(z^3 - 1)$ (cf. the proof of 17.20). The advantage here of cancellation by subtraction is that it enables us to build in the given values for the residues at the outset.

## Calculation of residues

To use Cauchy's residue theorem, we must be able to calculate residues, and finding residues by computing Laurent expansions is seldom appealing. In this section we derive formulae which enable residues to be worked out with a minimum of fuss.

**18.5 Classification of poles.** We already have a lot of information about the way poles arise and their relationship to zeros, from 15.3 and 17.11–17.13. Specifically, the function

$$f(z) = \frac{h(z)}{k(z)}$$

has a pole of order $m$ at $a$ if there exists $r > 0$ such that

(i) $h, k \in \mathrm{H}(\mathrm{D}(a; r))$,

(ii) $h(a) \neq 0$,

(iii) $k$ has a zero of order $m \geqslant 1$ at $a$, that is, it satisfies the equivalent conditions

(a) $k(a) = k'(a) = \cdots = k^{(m-1)}(a) = 0$, $k^{(m)}(a) \neq 0$,

(b) $k(z) = (z - a)^m g(z)$, where $g \in \mathrm{H}(\mathrm{D}(a; r))$ and $g(a) \neq 0$.

Assume that $f$ has a pole of order $m$ at $a$. We call the pole **simple** if $m = 1$ and **multiple** otherwise; **overt** (in the sense of being immediately visible) if $f(z)$ is expressed in the form $g(z)(z - a)^{-m}$, where $g \in \mathrm{H}(\mathrm{D}(a; r))$ for some $r > 0$ and $g(a) \neq 0$, and **covert** (that is, hidden) otherwise.

Whether a pole is overt or covert is a matter of how $f(z)$ is written. A covert pole of $h(z)/k(z)$ arises where a zero of $k$ is recognized through condition (a) and an overt pole if this is detected through condition (b). A covert pole can often be converted to an overt pole, but this is seldom prudent in residue calculations, especially if the pole is simple (see the examples in 18.10).

**18.6 Examples (overt and covert poles).**

- $1/((z-\mathrm{i})(z+\mathrm{i}))$ has overt simple poles at $\pm\mathrm{i}$; $1/(z^2+1)$ has covert simple poles at $\pm\mathrm{i}$.
- $\sec^2 z$ and $\tan^2 z$ both have covert double poles at $(2k+1)\pi/2$ $(k \in \mathbb{Z})$.

**18.7 The residue at a simple pole.** Let $f \in \mathrm{H}(\mathrm{D}(a;r))$ and assume that $f$ has a simple pole at $a$. We first observe that

(R) $$\operatorname{res}\{f(z);a\} = \lim_{z\to a}(z-a)f(z).$$

To prove (R), write $f(z)$ as $\sum_{n=-1}^{\infty} c_n(z-a)^n$ in $\mathrm{D}'(a;r)$. This implies that $\lim_{z\to a}(z-a)f(z) = c_{-1}$.

(1) **Overt simple pole** If $f(z) = g(z)/(z-a)$, where $g \in \mathrm{H}(\mathrm{D}(a;r))$ and $g(a) \neq 0$, then (R) implies that

$$\operatorname{res}\{f(z);a\} = g(a).$$

This is sometimes called the **cover-up rule.**

(2) **Covert simple pole** If $f(z) = h(z)/k(z)$, where $h,k \in \mathrm{H}(\mathrm{D}(a;r))$, $h(a) \neq 0$, $k(a) = 0$, and $k'(a) \neq 0$, then

$$\operatorname{res}\{f(z);a\} = \frac{h(a)}{k'(a)}.$$

We prove this as follows:

$$\begin{aligned}
\operatorname{res}\{f(z);a\} &= \lim_{z\to a}(z-a)\frac{h(z)}{k(z)} && \text{(by (R))}\\
&= h(a)\lim_{z\to a}\frac{z-a}{k(z)-k(a)} && \text{(by the algebra of limits)}\\
&= \frac{h(a)}{k'(a)}.
\end{aligned}$$

**18.8 The residue at a multiple pole.** Let $f$ have a pole of order $m > 1$ at $a$.

(1) **Overt multiple pole** Let $f(z) = g(z)/(z-a)^m$, where $g \in \mathrm{H}(\mathrm{D}(a;r))$ and $g(a) \neq 0$. Then

$$\operatorname{res}\{f(z);a\} = \frac{1}{(m-1)!}g^{(m-1)}(a).$$

This follows from Cauchy's formula for derivatives:

$$\begin{aligned} g^{(m-1)}(a) &= \frac{(m-1)!}{2\pi \mathrm{i}} \int_{\gamma(a;r/2)} \frac{g(z)}{(z-a)^m}\,\mathrm{d}z \\ &= \frac{(m-1)!}{2\pi \mathrm{i}} \int_{\gamma(a;r/2)} f(z)\,\mathrm{d}z \\ &= \operatorname{res}\{f(z); a\} \qquad \text{(by Lemma 18.1).} \end{aligned}$$

(2) **Covert multiple pole** No formula as neat as that for a covert simple pole exists. To find the residue, either convert to an overt pole or compute the Laurent coefficient $c_{-1}$. To find $c_{-1}$, write $w = (z-a)$ and expand in powers of $w$ for small $|w|$; use known expansions as far as possible. See 18.11 for examples.

**18.9 Residues of indeterminate forms.** We sometimes encounter functions $f(z) = h(z)/k(z)$ where $h$ and $k$ have zeros of orders $p$ and $q$ at $a$, with $q > p$. By 17.13, $f$ has a pole of order $m := q - p$ at $a$. To calculate the residue, we can compute the Laurent expansion about $a$ (in which a factor of $(z-a)^p$ will cancel). Alternatively, we can compute $\lim_{z\to a}((z-a)^m h(z))/k(z)$ with the aid of the complex form of L'Hôpital's rule (see Exercise 15.2).

**18.10 Examples (residues via the formulae in 18.7 and 18.8).**

**Tactical tip**

- ⊙ In applying the formulae 18.7(2) and 18.8(1) it is best to move into the numerator any factors in the denominator of $f(z)$ which do not contribute to there being a zero at $a$. For example, in computing the residue of $f(z) = 1/((z^3+1)\sin z)$ at $k\pi$ $(k \in \mathbb{Z})$ we would write $f(z)$ as $(z^3+1)^{-1}/\sin z$.

- $f(z) = 1/((2-z)(z^2+4))$ has simple poles at $z = 2$ (overt) and $z = \pm 2\mathrm{i}$ (covert):

$$\operatorname{res}\{f(z); 2\} = \operatorname{res}\left\{\frac{-(z^2+4)^{-1}}{z-2}; 2\right\} = -\frac{1}{8} \quad \text{(by 18.7(1)),}$$

$$\operatorname{res}\{f(z); \pm 2\mathrm{i}\} = \left[\frac{(2-z)^{-1}}{2z}\right]_{z=\pm 2\mathrm{i}} = \frac{1 \mp \mathrm{i}}{16} \quad \text{(by 18.7(2)).}$$

- $f(z) = 1/(1+z^4)$ has covert simple poles at the points $z_k = \mathrm{e}^{(2k+1)\pi \mathrm{i}/4}$ $(k = 0, 1, 2, 3)$. By 18.7(2),

$$\operatorname{res}\{f(z); z_k\} = \left[\frac{1}{4z^3}\right]_{z=z_k} = -\frac{1}{4}\mathrm{e}^{(2k+1)\pi \mathrm{i}/4},$$

since $z_k{}^4 = -1$. Factorization to convert to overt poles is not recommended here.

- $f(z) = 1/((z^2+1)\sin\pi z)$ has simple covert poles at $z = \pm\mathrm{i}$ and at $z = k$ $(k \in \mathbb{Z})$. By 18.7(2),

$$\operatorname{res}\{f(z);\pm\mathrm{i}\} = \left[\frac{\operatorname{cosec}\pi z}{2z}\right]_{z=\pm\mathrm{i}} = \frac{1}{\pm 2\mathrm{i}\sin\pi\mathrm{i}} = \mp\frac{1}{2\sinh\pi},$$
$$\operatorname{res}\{f(z);k\} = \left[\frac{(z^2+1)^{-1}}{\pi\cos\pi z}\right]_{z=k} = \frac{(-1)^k}{\pi(k^2+1)} \quad (k \in \mathbb{Z}).$$

- $f(z) = \mathrm{e}^{\mathrm{i}z}z^{-4}$ has an overt pole of order 4 at 0. By 18.8(1),

$$\operatorname{res}\{f(z);0\} = \frac{1}{3!}\left[\frac{\mathrm{d}^3}{\mathrm{d}z^3}\mathrm{e}^{\mathrm{i}z}\right]_{z=0} = -\frac{\mathrm{i}}{6}.$$

For an alternative method, via the Laurent expansion, see 18.11.

- $f(z) = (z+1)^{-2}(z^3-1)^{-1}$ has an overt double pole at $-1$ and covert simple poles at the cube roots of unity, $1$, $\omega$, $\omega^2$, where $\omega = \mathrm{e}^{2\pi\mathrm{i}/3}$. The simple poles are handled using 18.7(2). Let $\alpha$ be such that $\alpha^3 = 1$. Then

$$\operatorname{res}\{f(z);\alpha\} = \frac{(\alpha+1)^{-2}}{3\alpha^2}.$$

Hence $\operatorname{res}\{f(z);1\} = \frac{1}{12}$ and $\operatorname{res}\{f(z);\omega\} = \operatorname{res}\left\{f(z);\omega^2\right\} = \frac{1}{3}$ (remember that $\alpha^3 = 1$, $\alpha \neq 1$ implies $1+\alpha+\alpha^2 = 0$).

- $f(z) = (\pi\cot\pi z)/z^2$ has a covert simple pole at $z = k$ for each non-zero integer $k$, with

$$\operatorname{res}\left\{\frac{\pi\cot\pi z}{z^2};k\right\} = \frac{(\pi\cos\pi k/k^2)}{\pi\cos\pi k} = \frac{1}{k^2} \quad (k \in \mathbb{Z}\setminus\{0\}),$$

by 18.7(2). For the residue at the covert triple pole at 0; see 18.11.

**18.11 Examples (residues via the Laurent expansion).** For indeterminate forms, for multiple poles, or where the function is a product of several factors, it is often best to find the term in $(z-a)^{-1}$ in the Laurent expansion about the pole at $a$. Tactics for finding Laurent expansions were illustrated in 17.6.

- **A simple Laurent expansion** Consider again $f(z) = \mathrm{e}^{\mathrm{i}z}z^{-4}$. The Laurent expansion

$$\frac{\mathrm{e}^{\mathrm{i}z}}{z^4} = \frac{1}{z^4} + \frac{\mathrm{i}}{z^3} - \frac{1}{2!z^2} - \frac{\mathrm{i}}{3!z} + \ldots \quad (|z| > 0)$$

gives, by 18.2,

$$\operatorname{res}\{f(z); 0\} = -\frac{\mathrm{i}}{6}.$$

- **An indeterminate form** $z^{-3}\pi\cot\pi z$ has a covert pole at $z = 0$ of order 4. By 17.6,

$$\frac{\pi\cot\pi z}{z^3} = \frac{1}{z^3} - \frac{\pi^2}{3z} + \cdots = -\frac{\pi^2}{3} \quad \text{for small } |z|.$$

  Hence $\operatorname{res}\{f(z); 0\} = -\pi^2/3$.

- **Change of base point** Besides having an overt simple pole at zero, $f(z) = e^{\pi i z}/(z(4z-1)^3)$ has a triple pole at $z = 1/4$. Write $w = z - \frac{1}{4}$. Then, near $w = 0$,

$$\frac{e^{\pi\mathrm{i}(w+\frac{1}{4})}}{16w^3(1+4w)} = \frac{e^{\pi\mathrm{i}/4}}{16w^3}\left(1 + \pi\mathrm{i}w + \frac{(\pi\mathrm{i}w)^2}{2!} + \dots\right)\left(1 - 4w + 16w^2 + \dots\right).$$

  From this we can pick out the coefficient of the term in $w^{-1}$. Thus $\operatorname{res}\left\{f(z); \frac{1}{4}\right\}$ is $e^{\pi\mathrm{i}/4}(-4\pi\mathrm{i} + (\pi\mathrm{i})^2/2 + 16)/16$. We could use 18.8(2) instead, but this would not be any easier. The moral here is that calculating residues at multiple poles may be unavoidably messy.

**18.12 Integrals round the unit circle.** We shall use Cauchy's residue theorem to prove that

$$\int_0^{2\pi} \frac{1}{1 + 8\cos^2\theta}\,\mathrm{d}\theta = \frac{2\pi}{3}.$$

The method relies on converting the required integral into an integral round the unit circle. The circle is given parametrically by $z = \gamma(\theta) = e^{\mathrm{i}\theta}$ $(\theta \in [0, 2\pi])$. Then

$$\gamma'(\theta) = \mathrm{i}e^{\mathrm{i}\theta} = \mathrm{i}z,$$
$$\cos\theta = \tfrac{1}{2}(e^{\mathrm{i}\theta} + e^{-\mathrm{i}\theta}) = \tfrac{1}{2}(z + z^{-1}).$$

Hence

$$\begin{aligned}\int_0^{2\pi} \frac{1}{1 + 8\cos^2\theta}\,\mathrm{d}\theta &= \int_{\gamma(0;1)} \frac{1}{\mathrm{i}z}\,\frac{1}{\left(1 + 2(z^2 + 2 + z^{-2})\right)}\,\mathrm{d}z \\ &= \int_{\gamma(0;1)} \frac{z}{2z^4 + 5z^2 + 2}\,\mathrm{d}z \\ &= \int_{\gamma(0;1)} \frac{z}{(2z^2+1)(z^2+2)}\,\mathrm{d}z.\end{aligned}$$

The integrand has simple poles at $\pm\mathrm{i}/\sqrt{2}$ and $\pm\mathrm{i}\sqrt{2}$. The latter poles lie outside $\gamma(0;1)$ and do not contribute to the integral. So

$$\begin{aligned}\int_0^{2\pi} \frac{1}{1+8\cos^2\theta}\,\mathrm{d}\theta &= 2\pi\mathrm{i}\left(\operatorname{res}\left\{\frac{z}{(2z^2+1)(z^2+2)};\frac{\mathrm{i}}{\sqrt{2}}\right\}\right.\\ &\qquad\left.+\operatorname{res}\left\{\frac{z}{(2z^2+1)(z^2+2)};-\frac{\mathrm{i}}{\sqrt{2}}\right\}\right)\\ &= \left[\frac{2\pi\mathrm{i}}{4(z^2+2)}\right]_{z=\mathrm{i}/\sqrt{2}} + \left[\frac{2\pi\mathrm{i}}{4(z^2+2)}\right]_{z=-\mathrm{i}/\sqrt{2}}\end{aligned}$$

and this simplifies to $2\pi/3$.

This method can be used to evaluate various integrals of the general type

$$\int_0^{2\pi} F(\cos\theta,\sin\theta)\,\mathrm{d}\theta$$

by converting them to integrals round $\gamma(0;1)$ and applying Cauchy's residue theorem. Further examples can be found in Exercises 18.6 and 18.7.

**18.13 Evaluating complex integrals: taking stock.** Cauchy's residue theorem provides a natural way to attack a contour integral $\int_\gamma f(z)\,\mathrm{d}z$ when $f$ is holomorphic inside and on $\gamma$ except for at most finitely many poles inside $\gamma$ (and none on $\gamma^*$). En route to the residue theorem we proved several important results for evaluating integrals round contours:

- Cauchy's theorem (no poles),
- Cauchy's integral formula (a single pole, overt and simple),
- Cauchy's formula for derivatives (a single pole, overt and multiple; see 18.8(1)).

It is devious to cite Cauchy's residue theorem when one of these earlier results is applicable. The residue theorem comes into its own where there are covert poles. It also frees us from the restriction to just one singularity. (We can overcome this in simple cases by using partial fractions to split up a given integral. In the opposite direction, Exercise 19.3 shows that the existence of partial fraction decompositions can be derived from Cauchy's residue theorem.)

## Exercises

18.1 A function $f$ is holomorphic in $\mathbb{C}$ except for double poles at 1 and $-1$, of residues $a$ and $b$, respectively. It is also given that there exists a constant $K$ such that $|z^2 f(z)| \leqslant K$ for large $|z|$. Prove that $a + b = 0$. Find $f$ when $a = 1$ and $f(2\mathrm{i}) = f(-2\mathrm{i}) = 0$.

18.2 For each of the following functions, identify the type of pole (overt or covert, simple or multiple) that the function has at the indicated point $a$ and find the residue.

$$\text{(i)}\ \frac{z-3}{(z-1)(z-2)}\quad (a=2), \qquad \text{(ii)}\ \frac{\mathrm{e}^{\mathrm{i}z}}{z^6+1}\quad (a=\mathrm{e}^{\pi\mathrm{i}/6}),$$
$$\text{(iii)}\ \frac{z^2}{(z^2+1)^2}\quad (a=\mathrm{i}), \qquad \text{(iv)}\ \frac{1}{(z-1)^3(z-2)}\quad (a=1),$$
$$\text{(v)}\ \frac{\mathrm{e}^{\mathrm{i}z}-1}{z^2}\quad (a=0), \qquad \text{(vi)}\ \frac{\cot \pi z}{z^2+1}\quad (a=\mathrm{i}),$$
$$\text{(vii)}\ \frac{1}{z\sin^2 \pi z}\quad (a=-1), \qquad \text{(viii)}\ \frac{\mathrm{e}^{\mathrm{i}z}}{(1+z+z^2)}\quad (a=\omega).$$

18.3 Find the residues at the poles in $\mathbb{C}$ of the following functions (Exercises 17.8 and 17.9 sought identification of the singularities of each of these):

$$\text{(i)}\ \frac{1}{z^3(z^2+1)}, \qquad \text{(ii)}\ \frac{1}{z(3-z)}, \qquad \text{(iii)}\ \frac{1}{1-z+z^2},$$
$$\text{(iv)}\ \frac{1}{(z^2+z+1)^3}, \qquad \text{(v)}\ \frac{\mathrm{e}^{\mathrm{i}z}}{\cosh z}, \qquad \text{(vi)}\ \frac{\operatorname{cosec} \pi z}{z^2+1},$$
$$\text{(vii)}\ \frac{1}{z^4 \sin z}, \qquad \text{(viii)}\ \tan^2 z.$$

18.4 Compute

$$\text{(i)}\ \int_{\gamma(0;2)} \frac{1}{(z-1)^2(z^2+1)}\,\mathrm{d}z, \qquad \text{(ii)}\ \int_{\gamma(0;8)} (1+\mathrm{e}^z)^{-1}\,\mathrm{d}z.$$

18.5 Let $f(z) = \operatorname{cosec} \pi z$ and let $\{c_n\}$ and $\{d_n\}$ be the Laurent coefficients of $f$ in $\{\, z : 0 < |z| < 1 \,\}$ and $\{\, z : 1 < |z| < 2 \,\}$, respectively. Prove that $\frac{1}{2}(c_n - d_n)\pi = 1$ if $n$ is odd, and find the corresponding result when $n$ is even.

18.6 Prove the following by converting each integral into an integral round the unit circle and applying the residue theorem.

$$\text{(i)} \int_0^{2\pi} \frac{\sin^2\theta}{5+4\cos\theta}\,\mathrm{d}\theta = \frac{\pi}{4},$$

$$\text{(ii)} \int_0^{2\pi} \frac{1}{1+a\cos\theta}\,\mathrm{d}\theta = \frac{2\pi}{\sqrt{1-a^2}} \quad (-1<a<1),$$

$$\text{(iii)} \int_0^{2\pi} \frac{1}{(1+2a\cos\theta+a^2)^2}\,\mathrm{d}\theta = \frac{2\pi(1+a^2)}{(1-a^2)^3} \quad (-1<a<1).$$

18.7 Prove that, for any positive integer $n$,

$$\int_0^{2\pi} \sin^{2n}\theta\,\mathrm{d}\theta = \frac{2\pi}{4^n}\binom{2n}{n}.$$

The remaining exercises are theoretical ones, extending the theory of residues.

18.8 Let $f, g \in \mathrm{H}(\mathrm{D}(a;r))$ and assume that $f$ has a zero of order $m$ and $g$ a zero of order $m+1$ at $a$. Prove that

$$\operatorname{res}\left\{\frac{f(z)}{g(z)};a\right\} = (m+1)\frac{f^{(m)}(a)}{g^{(m+1)}(a)}.$$

18.9 (This exercise extends Theorem 15.13.) Let $f$ be holomorphic inside and on a positively oriented contour $\gamma$ except at a finite number of poles inside $\gamma$. Assume that $f(z) \neq 0$ for $z \in \gamma^*$ and that $f$ has $N$ zeros and $P$ poles inside $\gamma$, a zero or pole of order $m$ being counted $m$ times. Prove that

$$\frac{1}{2\pi\mathrm{i}}\int_\gamma \frac{f'(z)}{f(z)}\,\mathrm{d}z = N - P.$$

18.10 Let $f$ be holomorphic inside and on $\gamma(a;r)$ and assume that $f(z) \neq 0$ for $z \in \gamma(a;r)^*$. Find, in terms of the zeros of $f$,

$$\frac{1}{2\pi\mathrm{i}}\int_{\gamma(a;r)} \frac{f'(z)}{f(z)z^m}\,\mathrm{d}z \quad (m = 1,2,3,\ldots).$$

(Theorem 15.13 treats the case $m = 0$.)

# 19 A technical toolkit for contour integration

Cauchy's residue theorem is a powerul tool, as the applications in later chapters will amply demonstrate. The theorem relates the integral of a meromorphic function $f$ round a positively oriented contour $\gamma$ to the residues at the poles of $f$ inside $\gamma$. We already have good techniques for calculating residues, but before we can exploit the residue theorem to the full we need further information about integrals along paths.

## Evaluating real integrals by contour integration

**19.1 An introductory example.** Suppose we wish to evaluate the real integral

$$I := \int_0^\infty \frac{1}{1+x^4}\,\mathrm{d}x.$$

(Here the integral is interpreted to mean $\lim_{R\to\infty}\int_0^R (1+x^4)^{-1}\,\mathrm{d}x$. For a brief discussion of the definition of integrals over $[0,\infty)$ and over $(-\infty,\infty)$ see 19.14.) Because the integrand is an even function,

$$2\int_0^R \frac{1}{1+x^4}\,\mathrm{d}x = \int_{-R}^R \frac{1}{1+x^4}\,\mathrm{d}x.$$

Let $\gamma$ be the semicircular contour $\Gamma(0;R)$ shown in Fig. 19.1. We have

$$\int_\gamma \frac{1}{1+z^4}\,\mathrm{d}z = \int_{-R}^R \frac{1}{1+x^4}\,\mathrm{d}x + \int_0^\pi \frac{R\mathrm{i}\mathrm{e}^{\mathrm{i}\theta}}{1+R^4\mathrm{e}^{4\mathrm{i}\theta}}\,\mathrm{d}\theta.$$

As $R\to\infty$, the first term on the right-hand side tends to $2I$, while the second, as we showed in 10.11, tends to zero.

The value of $\int_\gamma (1+z^4)^{-1}\,\mathrm{d}z$ can be computed by Cauchy's residue theorem: $(1+z^4)^{-1}$ is holomorphic inside and on $\gamma$ except for covert simple poles at the zeros of $(1+z^4)$. These zeros are at $\mathrm{e}^{(2k+1)\pi\mathrm{i}/4}$ $(k=0,1,2,3)$. Only those at

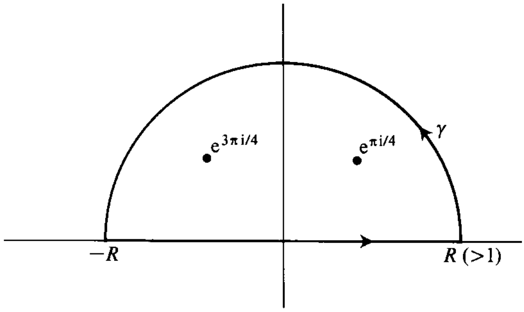



Figure 19.1 Contour $\Gamma(0;R)$ for Example 19.1

$z_1 := e^{i\pi/4} = (1+i)/\sqrt{2}$ and $z_2 := e^{3\pi i/4} = (-1+i)/\sqrt{2}$ lie inside $\gamma$; the other two zeros do not concern us (see Fig. 19.1). We have, from 18.10,

$$\operatorname{res}\left\{(1+z^4)^{-1}; z_k\right\} = -\frac{1}{4}z_k.$$

Hence, by Cauchy's residue theorem,

$$\int_\gamma \frac{1}{1+z^4}\,dz = -\frac{2\pi i}{4}\left(e^{\pi i/4} + e^{3\pi i/4}\right) = -\frac{2\pi i}{4}\left(\frac{1+i}{\sqrt{2}} + \frac{-1+i}{\sqrt{2}}\right) = \frac{\pi}{\sqrt{2}}.$$

We conclude that

$$\int_0^\infty \frac{1}{1+x^4}\,dx = \frac{\pi}{2\sqrt{2}}.$$

**Tactical tip**

⊙ This is an opportune moment at which to issue a reminder that, when Cauchy's residue theorem is used, only the residues at the poles *inside* the contour contribute to the integral.

**19.2 Towards a general strategy.** We have now seen two examples of the use of Cauchy's residue theorem for the evaluation of real definite integrals. In 18.12 we were able to convert the required integral into one round the unit circle to which the residue theorem was directly applicable. In Example 19.1 the situation was more complicated. We proceeded in stages:

(1) The required integral was viewed as the limit of an integral along the segment $[-R, R]$ of the real axis.

(2) By joining this segment with the arc $\Gamma_R$, we formed the contour $\gamma$.

(3) We found a function $f(z) = (1+z^4)^{-1}$ coinciding with the required integrand $(1+x^4)^{-1}$ when $z$ is the real variable $x$, with $f$ holomorphic inside and on $\gamma$ except for finitely many poles inside $\gamma$.

(4) We estimated the (unknown) integral of $f(z)$ along the semicircular arc $\Gamma_R$ and discovered—providentially—that it tended to zero as $R \to \infty$.

(5) We calculated the residues at the poles of $f$ inside $\gamma$ and applied Cauchy's residue theorem.

(6) Finally, we let $R \to \infty$.

Key elements in this strategy are the following:

- we relate the required integral $I$, or an approximation to it, to some contour integral $\int_\gamma f(z)\,\mathrm{d}z$;
- we must be able to apply Cauchy's residue theorem, so that $f$ must have at most finitely many poles inside the contour $\gamma$ and none on it;
- the contour $\gamma$ is chosen so that the integral of $f$ along each portion of it either contributes to the integral we want, or can be handled by estimation or in some other way.

The power of the method rests in its versatility and this makes it difficult to lay down hard and fast rules on how to use it. Accordingly, it is best learned through examples. In Chapter 20 we carry through our general strategy for integrals of various types. A number of technical tools are needed repeatedly, especially those used for estimating integrals or finding their limits. The rest of this short chapter collects these tools together, for ease of reference.

## Inequalities and limits

**19.3 Basic inequalities.** The key result for finding an upper bound for the modulus of an integral is the Estimation theorem (10.10): given a path $\gamma$ with parameter interval $[\alpha, \beta]$ and a continuous function $f$ on $\gamma^*$,

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant \int_\alpha^\beta |f(\gamma(t))\gamma'(t)|\,\mathrm{d}t \leqslant \sup_{z\in\gamma^*} |f(z)| \times \mathrm{length}(\gamma).$$

To obtain an upper bound for $|f(\gamma(t))\gamma'(t)|$, our first line of attack is to use one or more of the following, which are true for complex numbers $z_1, z_2, \ldots$:

(1) $|z_1 + z_2| \leqslant |z_1| + |z_2|$ (1.9(2));

(2) $|z_1 + z_2| \geqslant ||z_1| - |z_2||$ (1.9(3));

(3) $|z_1| \leqslant |z_2| \iff 1/|z_1| \geqslant 1/|z_2|$ (for $z_1, z_2 \neq 0$);

(4) $1/|z_1 + z_2| \leqslant 1/||z_1| - |z_2||$ (combine (2) and (3)).

In (2), the exterior modulus signs ensure that we get a meaningful inequality irrespective of which of $|z_1|$ and $|z_2|$ is the larger. Inequalities (1) and (4) are used, respectively, for obtaining bounds on numerators and denominators in fractional expressions. In (4), note particularly the minus sign. If in doubt when handling $1/|z_1 + z_2|$, remember that to make this bigger we must make the denominator smaller, and that this is not achieved by replacing $|z_1 + z_2|$ by $|z_1| + |z_2|$. Iterating, we get the following, for $n \geqslant 3$:

(5) $|z_1 + z_2 + \cdots + z_n| \leqslant |z_1| + \cdots + |z_n|$

(6) $|z_1 + z_2 + \cdots + z_n| \geqslant |z_1| - |z_2| - \cdots - |z_n|$ if $|z_1| \geqslant |z_2| + \cdots + |z_n|$.

This is an opportune time to issue a reminder that inequalitites must be between *real* numbers; see 1.9. Omission of moduli in the Estimation theorem or in (1)–(6) causes havoc.

**19.4 Bounds on exponential factors.** We often need a bound on an exponentiai $e^{iz}$ on an arc. On $|z| = r$, we have

$$\left|e^{iz}\right| = \left|e^{ire^{i\theta}}\right| = \left|e^{ir(\cos\theta + i\sin\theta)}\right| = e^{-r\sin\theta} \quad (z = re^{i\theta}).$$

Now consider $r = R$, where $R$ is large. On the semicircular arc $\Gamma_R$ in the upper half-plane we have $0 \leqslant \theta \leqslant \pi$, so $\sin\theta \geqslant 0$. This bound is often adequate, but is in fact very crude when $0 < \theta < \pi$ and $R$ is very large. For example, when $\theta = \pi/2$, we have $e^{-R\sin\theta} = e^{-R}$, which is very small by comparison with 1 when $R$ is large. In situations where the condition $\sin\theta \geqslant 0$ gives too crude a bound on an expression involving $\sin\theta$, we may get an improved bound from the following result. See 20.14 and 20.16 for examples using this technique.

**19.5 Jordan's inequality.** $\dfrac{2}{\pi} \leqslant \dfrac{\sin\theta}{\theta} \leqslant 1$ for $0 < \theta \leqslant \frac{1}{2}\pi$.

**Proof** It will be sufficient to prove that $\sin\theta/\theta$ decreases as $\theta$ increases for $\theta \in (0, \frac{1}{2}\pi]$. This is the case if

$$\frac{d}{d\theta}\left(\frac{\sin\theta}{\theta}\right) \leqslant 0 \quad \text{for } 0 < \theta \leqslant \tfrac{1}{2}\pi.$$

But

$$\frac{d}{d\theta}\left(\frac{\sin\theta}{\theta}\right) = \frac{\theta\cos\theta - \sin\theta}{\theta^2} \leqslant 0 \quad \text{whenever } \theta\cos\theta - \sin\theta \leqslant 0.$$

Since $[\theta\cos\theta - \sin\theta]_{\theta=0} = 0$, it is now enough to note that, on $(0, \frac{1}{2}\pi]$, the function $\theta\cos\theta - \sin\theta$ has a non-positive derivative and so decreases as $\theta$ increases. $\square$

**19.6 Basic limits.** The following real limits are probably already familiar. Since we shall use them frequently, we include proofs.

(1) For any constant $k > 0$, $x^k e^{-x} \to 0$ as $x \to \infty$ $(x \in \mathbb{R})$.

(2) For any constant $k > 0$,

$$x^{-k}\log x \to 0 \text{ as } x \to \infty \quad \text{and} \quad x^k \log x \to 0 \text{ as } x \to 0 \quad (x \in \mathbb{R},\ x > 0).$$

**Proof** (1) For $x > 0$,

$$0 < x^k e^{-x} = \frac{x^k}{1 + x + \cdots + x^n/n! + \ldots} < n!x^{k-n}.$$

This is true for any $n$. Since $n$ can be chosen greater than $k$, the result follows.

(2) To obtain the first result, put $x = e^{y/k}$ and note that

$$x^{-k}\log x = (y/k)e^{-y},$$

and, by (1), this tends to zero as $y \to \infty$, and hence as $x \to \infty$. The second limit follows from the first on replacing $x$ by $x^{-1}$. $\square$

## Estimation techniques

**19.7 Estimation of integrals round large circular arcs.** We frequently want to create a contour by joining the endpoints of a path by a circular arc of radius $R$. The integral along this arc will rarely be explicitly computable and we shall hope to show that it tends to zero as $R \to \infty$. Let $f$ be continuous on $\gamma^*$, where $\gamma(\theta) = Re^{i\theta}$ $(\theta_1 \leqslant \theta \leqslant \theta_2)$. We have $\gamma'(\theta) = Rie^{i\theta}$, and the Estimation theorem gives

$$\left|\int_\gamma f(z)\,dz\right| \leqslant \int_{\theta_1}^{\theta_2} \left|f(Re^{i\theta})\right| R\,d\theta.$$

Note particularly the factor $R$ coming from $\gamma'(\theta)$.

- Take $f(z) = (z^2 + z + 1)^{-2}$. By 19.3(3) and (5),

$$\left| \int_\gamma f(z)\, \mathrm{d}z \right| \leqslant \int_{\theta_1}^{\theta_2} \frac{1}{R^2 - R - 1} R\, \mathrm{d}\theta \quad (\text{if } R^2 - R - 1 > 0)$$
$$= \mathbf{O}(R^{-1}).$$

- Let $f(z) = \mathrm{e}^{\mathrm{i}z} z^{-k}$ $(k > 1)$ and $\gamma = \Gamma_R$ (so $\theta_1 = 0$ and $\theta_2 = \pi$). Then

$$|f(R\mathrm{e}^{\mathrm{i}\theta})| = \frac{|\mathrm{e}^{\mathrm{i}(R\cos\theta + \mathrm{i}R\sin\theta)}|}{|R^k \mathrm{e}^{\mathrm{i}k\theta}|} = \frac{\mathrm{e}^{-R\sin\theta}}{R^k}.$$

On $[0, \pi]$, we have $\sin\theta \geqslant 0$, so $\mathrm{e}^{-R\sin\theta} \leqslant 1$. Hence

$$\left| \int_{\Gamma_R} f(z)\, \mathrm{d}z \right| \leqslant \int_0^\pi R^{1-k}\, \mathrm{d}\theta \to 0 \text{ as } R \to \infty.$$

- Let $f(z) = \mathrm{e}^{\mathrm{i}z} z^{-k}$ $(0 < k \leqslant 1)$ and again take $\gamma = \Gamma_R$. Using the fact that $\sin\theta \leqslant 1$ as above, we get

$$\left| \int_{\gamma_R} f(z)\, \mathrm{d}z \right| \leqslant \int_0^\pi R^{1-k}\, \mathrm{d}\theta.$$

The right-hand side does not tend to zero as $R \to \infty$. So we seek a tighter bound on $\sin\theta$ is available from Jordan's inequality:

$$\left| \int_{\gamma_R} f(z)\, \mathrm{d}z \right| \leqslant \int_0^\pi \mathrm{e}^{-R\sin\theta} R^{1-k}\, \mathrm{d}\theta$$
$$= 2\int_0^{\pi/2} \mathrm{e}^{-R\sin\theta} R^{1-k}\, \mathrm{d}\theta$$
$$\leqslant 2R^{1-k} \int_0^{\pi/2} \mathrm{e}^{-2R\theta/\pi}\, \mathrm{d}\theta$$
$$= \pi R^{-k}(1 - \mathrm{e}^{-R}) \leqslant \pi R^{-k},$$

and this does tend to zero as $R \to \infty$.

**Tactical tips**

⊙ In this example Jordan's inequality produces a life-saving factor of $R^{-1}$. The inequality should *only* be used where a rougher estimate, as in 19.4, is no help.

⊙ Note how it was first necessary to change the range of integration to $[0, \pi/2]$; Jordan's inequality is not valid on $[0, \pi]$.

**19.8 The need for indented contours.** Suppose we want to integrate a function $f$ round a contour $\gamma$. We may be thwarted because there is on $\gamma^*$ either a pole of $f$ or (in the case that $f$ is a branch of a multifunction) a branch point of $f$. A possible remedy is to 'walk round' the offending point, following a circular arc of small radius $\varepsilon$, as illustrated in Fig. 19.2. Our hope will be that we can compute the limit of the integral round an indentation as its radius shrinks to 0. In 19.9–19.12 we indicate when this is possible and when it is not.

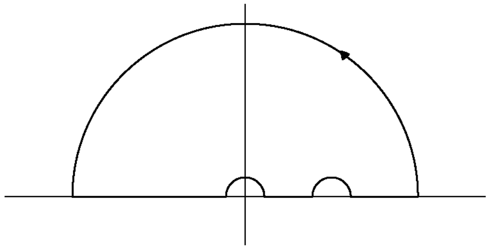


**Figure 19.2** A semicircle with indentations

Cauchy's residue theorem tells us that $\int_{\gamma(a;\varepsilon)} f(z)\,\mathrm{d}z = 2\pi \mathrm{i} b$ when $f$ has a simple pole at $a$ of residue $b$. What the Indentation lemma tells us is that *in the limit as* $\varepsilon \to 0$, the integral of $f$ along a portion of $\gamma(a;\varepsilon)$ is $2\pi \mathrm{i} b$ times the fraction of $\gamma(a;\varepsilon)^*$ traversed.

**19.9 Indentation lemma (for a simple pole).** Let $f \in \mathrm{H}(\mathrm{D}'(a;r))$ and let $f$ have a simple pole of residue $b$ at $a$. Let an indentation round $a$ be given by $\gamma_\varepsilon(\theta) = a + \varepsilon \mathrm{e}^{\mathrm{i}\theta} \quad (\theta \in [\theta_1, \theta_2])$, where $0 < \varepsilon < r$ and $0 \leqslant \theta_1 < \theta_2 \leqslant 2\pi$. Then

$$\lim_{\varepsilon \to 0} \int_{\gamma_\varepsilon} f(z)\,\mathrm{d}z = \mathrm{i}b(\theta_2 - \theta_1).$$

**Proof** From 18.7, $b = \lim_{z\to a}(z-a)f(z)$. Given $\eta > 0$, there exists $\delta > 0$ such that $|(z-a)f(z) - b| < \eta$ whenever $0 < |z-a| < \delta$. Let $0 < \varepsilon < \min\{r, \delta\}$. When $z = \gamma_\varepsilon(\theta)$ we have $\gamma_\varepsilon'(\theta) = \varepsilon \mathrm{i}\mathrm{e}^{\mathrm{i}\theta} = \mathrm{i}(z-a)$ and so

$$\begin{aligned}\left|\int_{\gamma_\varepsilon} f(z)\,\mathrm{d}z - \mathrm{i}b(\theta_2-\theta_1)\right| &= \left|\int_{\theta_1}^{\theta_2} \big(f(\gamma_\varepsilon(\theta))\gamma_\varepsilon'(\theta) - \mathrm{i}b\big)\,\mathrm{d}\theta\right| \\ &= \left|\int_{\theta_1}^{\theta_2} g(\gamma_\varepsilon(\theta))\,\mathrm{d}\theta\right| \\ &< \eta(\theta_2-\theta_1). \qquad \square\end{aligned}$$

**19.10 Indentation at a multiple pole.** Not allowed! The formula in the Indentation lemma does *not* apply when the pole at $a$ is not simple. Indeed, $\lim_{\varepsilon\to 0}\int_{\gamma_\varepsilon} f(z)\,\mathrm{d}z$ does not even exist, essentially because $|f(z)|$ 'blows up' too fast as $z$ approaches $a$.

**19.11 Integration in a cut plane.** We would like to be able to apply Cauchy's theorem or the residue theorem to integrals $\int_\gamma f(z)\,\mathrm{d}z$, where $f$ is, or $f$ incorporates, a holomorphic branch of a multifunction. Branches can be specified directly in terms of argument restrictions, without the device of cutting the plane. But cuts do make it transparent that the contour $\gamma$ is allowable provided it lies in the cut plane (and so crosses no cut).

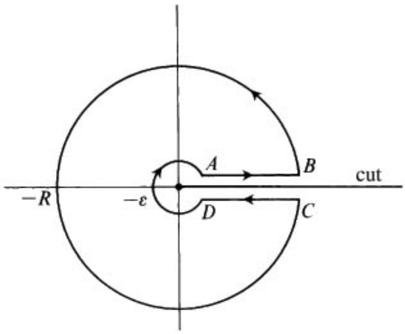



**Figure 19.3** A keyhole contour

In Chapter 9 we were somewhat schizophrenic about the status of points on cuts. Defining the value of a branch on one edge of a cut but not the other prevents branches being 2-valued on the cut, but does seem a little arbitrary. However, in certain applications, it is tidier to keep both edge-values in play at the points of a cut. Consider the following example. In the plane cut along $[0,\infty)$, take $f$ to be the branch of the logarithm given by

$$f(z) = \log r + \mathrm{i}\theta \quad (0 \neq z = r\mathrm{e}^{\mathrm{i}\theta},\ 0 \leqslant \theta < 2\pi).$$

We cannot integrate $f$ around $\gamma(0;R)$, since this crosses the cut. Suppose that we integrate instead round the keyhole contour $\gamma$ shown in Fig. 19.3. In the cut plane we can take the horizontal lines to be at an arbitrarily small distance $\delta > 0$ from the real axis. Then $f$ is holomorphic inside and on $\gamma$. Also,

$$\lim_{\delta\to 0}\int_{AB} f(z)\,\mathrm{d}z = \int_\varepsilon^R \lim_{\delta\downarrow 0} f(x+\delta\mathrm{i})\,\mathrm{d}x = \int_\varepsilon^R \log t\,\mathrm{d}t$$

and

$$\lim_{\delta \to 0} \int_{CD} f(z)\,\mathrm{d}z = \int_0^R \lim_{\delta \downarrow 0} f(x - \delta \mathrm{i})\,\mathrm{d}x = \int_0^R (\log t + 2\pi\mathrm{i})\,\mathrm{d}t.$$

The verification of these claims is technical, but not difficult. Rather than go through this sort of limiting process every time we integrate in a cut plane—and we do so quite frequently in the following chapters—we shall allow ourselves the liberty of integrating along the edge or edges of a cut, using the obvious edge-values for the integrand.

**19.12 Indentation at a branch point.** Suppose $a$ is a branch point of $f$. Then $f$ is not holomorphic in any punctured disc centre $a$. Therefore it does not have an isolated singularity at $a$. In particular, $a$ cannot be a pole of $f$ and the formula in the Indentation lemma does not apply. Usually the basic inequalities in 19.3 will show that $\int_{\gamma_\varepsilon} f(z)\,\mathrm{d}z \to 0$ as $\varepsilon \to 0$.

We conclude this section with a boundedness result we shall need when evaluating certain infinite sums in Chapter 20. It will enable us to show that the integrals of certain functions of the form $\varphi(z)\cot \pi z$ and $\varphi(z)\operatorname{cosec} \pi z$ tend to zero. Notice that the squares $S_N$ in the lemma avoid the poles of both $\cot \pi z$ and $\operatorname{cosec} \pi z$ (which are at each $k \in \mathbb{Z}$).

**19.13 Boundedness lemma for cot and cosec.** Let $S_N$ be the square with vertices at $(\pm 1 \pm \mathrm{i})(N + \frac{1}{2})$ $(N = 1, 2, \dots)$.

(1) There exists a constant $C$ such that $|\cot \pi z| \leqslant C$ for all $z \in S_N$.

(2) There exists a constant $K$ such that $|\operatorname{cosec} \pi z| \leqslant K$ for all $z \in S_N$.

**Proof** (1) On the horizontal sides of $S_N$, we have $z = x \pm \mathrm{i}(N + \frac{1}{2})$ and so, by basic inequalities and the formulae in 7.6 and 7.7,

$$\begin{aligned} |\cot \pi z| &= \left| \frac{\mathrm{e}^{\mathrm{i}\pi(x \pm \mathrm{i}(N+\frac{1}{2}))} + \mathrm{e}^{-\mathrm{i}\pi(x \pm \mathrm{i}(N+\frac{1}{2}))}}{\mathrm{e}^{\mathrm{i}\pi(x \pm \mathrm{i}(N+\frac{1}{2}))} - \mathrm{e}^{-\mathrm{i}\pi(x \pm \mathrm{i}(N+\frac{1}{2}))}} \right| \\ &= \frac{\mathrm{e}^{\pi(N+\frac{1}{2})} + \mathrm{e}^{-\pi(N+\frac{1}{2})}}{\mathrm{e}^{\pi(N+\frac{1}{2})} - \mathrm{e}^{-\pi(N+\frac{1}{2})}} \\ &\leqslant \coth \pi(N + \tfrac{1}{2}) \\ &\leqslant \coth \tfrac{3}{2}\pi. \end{aligned}$$

(For the final inequality, note that $\coth t$ is a decreasing function of $t$ for $t \geqslant 0$.)

On the vertical sides of $S_N$, we have $z = \pm(N + \frac{1}{2}) + \mathrm{i}y$ and so, by the trigonometric addition formulae,

$$|\cot \pi z| = |\tan \mathrm{i}\pi y| = |\tanh \pi y| \leqslant 1.$$

(2) is left as an exercise. □

**Tactical tip**

⊙ To see why trigonometric formulae have been used for the vertical sides of the square and exponential formulae for the horizontal sides, experiment with using these the opposite way round.

## Improper and principal-value integrals

Many of the integrals we shall evaluate by contour integration will be of the form

$$\int_0^\infty f(x)\,\mathrm{d}x \quad \text{or} \quad \int_{-\infty}^\infty f(x)\,\mathrm{d}x.$$

This section clarifies how we interpret such integrals. We assume that $f$ is either real- or complex-valued.

**19.14 Definitions (improper and principal-value integrals).** For a function defined on a closed bounded subinterval of $\mathbb{R}$, different ways of defining the integral lead to the same result so long as $f$ is minimally well behaved. Certainly this is the case for the functions we need to consider, which are, at the very least, piecewise continuous. Therefore, in the definitions below the term 'integrable' can be taken as referring to the reader's preferred theory of integration.

(1) Suppose that $f$ is integrable on every interval $[0, R]$ $(0 < R < \infty)$. The **improper integral of** $f$ **over** $[0, \infty)$ is defined to be

$$\lim_{R\to\infty} \int_0^R f(x)\,\mathrm{d}x,$$

if this limit exists.

(2) Suppose that $f$ is integrable on every bounded closed subinterval of $\mathbb{R}$. The **improper integral of** $f$ **over** $(-\infty, \infty)$ is defined to be

$$\lim_{R,S\to\infty} \int_{-S}^R f(x)\,\mathrm{d}x,$$

if this limit exists; here $R$ and $S$ tend to $\infty$ independently.

(3) Suppose that $f$ is integrable on every subinterval $[-R, R]$ $(R > 0)$ of $\mathbb{R}$. The **principal-value integral of $f$ over** $(-\infty, \infty)$ is defined to be

$$\mathrm{PV}\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x := \lim_{R\to\infty}\int_{-R}^{R} f(x)\,\mathrm{d}x.$$

**19.15 Comparison and reconciliation of the definitions in 19.14.** Principal-value and improper integrals arise naturally from limits of contour integrals. Consider, for example, the integral of $f(z)$ round the contours $\gamma_1$ and $\gamma_2$ shown in Fig. 19.4: $\int_{\gamma_1} f(z)\,\mathrm{d}z$ incorporates $\int_{-R}^{R} f(x)\,\mathrm{d}x$ and $\int_{\gamma_2} f(x)\,\mathrm{d}x$ incorporates $\int_{-S}^{R} f(x)\,\mathrm{d}x$. A semicircular indentation round a point $a \in \mathbb{R}$ can also give rise to a PV-style integral (where an interval $(a-\varepsilon, a+\varepsilon)$ is excluded from the path and $\varepsilon$ allowed to tend to zero).

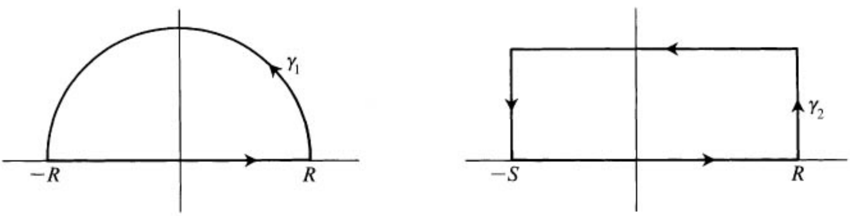



**Figure 19.4** Sample contours

Trivially, if the improper integral of $f$ over $(-\infty, \infty)$ exists, then so does the principal-value integral, and the two coincide. The converse fails: $\int_{-S}^{R} x\,\mathrm{d}x = \frac{1}{2}(R^2 - S^2)$, which does not tend to a limit as $R, S \to \infty$ independently; on the other hand, $\int_{-R}^{R} x\,\mathrm{d}x = 0$ for all $R$, so $\mathrm{PV}\int_{-\infty}^{\infty} x\,\mathrm{d}x$ exists and equals zero.

In Riemann integration, the integral of $f$ over $[0, \infty)$ or $(-\infty, \infty)$ is, *by definition*, the improper integral over the interval in question. In Lebesgue integration, integrals over infinite intervals are not *defined* via improper integrals. However the Dominated convergence theorem shows that if $f$ has a Lebesgue integral over $[0, \infty)$ then its improper integral over $[0, \infty)$ exists, and equals the Lebesgue integral, and similarly for $(-\infty, \infty)$. The converse fails: the classic example is provided by $f(x) = x^{-1}\sin x$. This is not Lebesgue integrable over

$[0, \infty)$, nor does the improper Riemann integral of $|f|$ exist. However $f$ does have an improper integral over $[0, \infty)$. (See, for example, [**6**] or [**2**].)

We henceforth adopt the following conventions:

- $\int_0^\infty f(x)\,dx$ means either the improper (Riemann) integral [or, for Lebesgue integral *afficiandos*, the Lebesgue integral when this exists].
- $\int_{-\infty}^\infty f(x)\,dx$ means the PV-integral when this exists but the improper integral does not, and the improper (Riemann) integral otherwise. [The latter may be interpreted as a Lebesgue integral when $f$ is Lebesgue integrable on $\mathbb{R}$.]

We note that a sufficient condition for $f$ to have [a Lebesgue integral and hence] an improper integral on $(-\infty, \infty)$ is

(i) $f$ is integrable on any closed bounded subinterval of $\mathbb{R}$, and
(ii) $f(x) = \mathbf{O}(|x|^{-p})$ for large $|x|$, for some constant $p > 1$.

## Exercises

**Exercises from the text.** Prove the Boundedness lemma for cosec (19.13(2)).

19.1 Evaluate the following limits:

(i) $\displaystyle\lim_{R\to\infty}\int_{\Gamma_R} \frac{e^{iz}}{z^4+z^3+z^2+z+1}\,dz$,

(ii) $\displaystyle\lim_{R\to\infty}\int_{\Gamma_R} \left(\frac{z-1}{z^2+1}\right)e^{iz}\,dz$,

(iii) $\displaystyle\lim_{R\to\infty}\int_{\gamma(0;R)} \frac{p(z)}{q(z)}\,dz$, where $p$ and $q$ are polynomials and $\deg p < \deg q - 1$.

19.2 Evaluate the following limits:

$$\text{(i) } \lim_{\varepsilon\to 0}\int_{\Gamma_\varepsilon} \operatorname{cosec}\pi z\,dz, \qquad \text{(ii) } \lim_{\varepsilon\to 0}\int_{-\Gamma_\varepsilon} \frac{e^{iz}}{z(z^2+1)}\,dz.$$

19.3 Suppose that $p$ and $q$ are polynomials of degrees $m$ and $n$ respectively, where $n \geqslant m+1$, and that $q$ has simple zeros at $b_1, \ldots, b_n$. By integrating $f(w) = p(w)/(q(w)(w-z))$ round $\gamma(0; R)$ for large $R$, obtain the partial fraction decomposition

$$\frac{p(z)}{q(z)} = \sum_{k=1}^{n} \frac{p(b_k)}{q'(b_k)}(z-b_k)^{-1}.$$

Hence decompose $(1 - z^2)/(1 + z^4)$ into partial fractions.

19.4 Suppose that $f$ is holomorphic inside and on $\gamma(0;1)$. By integrating round the contour shown in Fig. 19.3, prove that

$$\int_0^1 f(x)\,\mathrm{d}x = \frac{1}{2\pi\mathrm{i}}\int_{\gamma(0;1)} f(z)(\log z - \mathrm{i}\pi)\,\mathrm{d}z,$$

where $\log z$ denotes the branch of the logarithm whose imaginary part lies between 0 and $2\pi$. Deduce that

$$\left|\int_0^1 f(x)\,\mathrm{d}x\right| \leqslant \frac{1}{2}\int_0^{2\pi} \left|f(\mathrm{e}^{\mathrm{i}\theta})\right|\mathrm{d}\theta.$$

# 20 Applications of contour integration

This chapter is devoted to applications of Cauchy's residue theorem to the evaluation of definite integrals and the summation of series. The method will handle quite baroque examples. If we seem to have included some examples of this sort, especially in the final section (advanced track), it is because these give valuable technical experience. Very few of the integrals we consider can be evaluated by more elementary methods such as substitution.

Further methods and examples can be found in Chapters 21 and 22. In particular, integrals yielding the characteristic functions of some well-known probability distributions are considered in Chapter 22. Recall also the use of the residue theorem to compute trigonometric integrals over $[0, 2\pi]$; see 18.12 and the exercises for Chapter 19.

## Integrals of rational functions

Our first example uses the same technique as we employed in 19.1.

**20.1 Example.** To evaluate $\displaystyle\int_0^\infty \frac{1}{(x^2+1)^2(x^2+4)}\,\mathrm{d}x.$

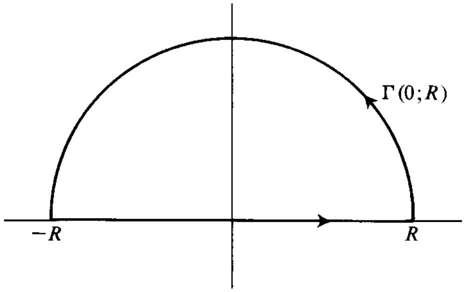



**Figure 20.1** Contour $\Gamma(0; R)$ for Examples 20.1 and 20.3

**Solution** We integrate $f(z) = 1/((z^2+1)^2(z^2+4))$ round the contour $\gamma = \Gamma(0; R)$ in Fig. 20.1 (choosing $R > 2$). The function $f$ is holomorphic inside and on $\gamma$, except for a double pole at $\mathrm{i}$ and a simple pole at $2\mathrm{i}$. By Cauchy's residue theorem,

$$\int_{-R}^{R} f(z)\,\mathrm{d}z + \int_{\Gamma_R} f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\big(\mathrm{res}\,\{f(z);\mathrm{i}\} + \mathrm{res}\,\{f(z);2\mathrm{i}\}\big).$$

By 18.8(1) and 18.7(2),

$$\mathrm{res}\,\{f(z);\mathrm{i}\} = \left[\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{(z+\mathrm{i})^2(z^2+4)}\right]_{z=\mathrm{i}} = \left[\frac{-2z(z+\mathrm{i}) - 2(z^2+4)}{(z+\mathrm{i})^3(z^2+4)^2}\right]_{z=\mathrm{i}} = -\frac{\mathrm{i}}{36},$$

$$\mathrm{res}\,\{f(z);2\mathrm{i}\} = \left[\frac{1}{(z^2+1)^2(2z)}\right]_{z=2\mathrm{i}} = -\frac{\mathrm{i}}{18}.$$

Also (see 19.3),

$$\left|\int_{\Gamma_R} f(z)\,\mathrm{d}z\right| \leqslant \int_0^{\pi} \frac{1}{(R^2-1)^2(R^2-4)} R\,\mathrm{d}\theta = \mathbf{O}(R^{-5})$$

and

$$\int_{-R}^{R} f(x)\,\mathrm{d}x = 2\int_0^R \frac{1}{(x^2+1)^2(x^2+4)}\,\mathrm{d}x.$$

Hence, letting $R \to \infty$,

$$\int_0^{\infty} \frac{1}{(x^2+1)^2(x^2+4)}\,\mathrm{d}x = \frac{\pi}{6}.$$

**Tactical tip**

⊙ We would not have been able to obtain the integral of $f(x)$ over $[0,\infty)$ using the contour $\Gamma(0;R)$ if $f$ had not been an even function on $\mathbb{R}$, that is, such that $f(x) = f(-x)$ for $x$ real.

**20.2 Example.** To evaluate $I := \displaystyle\int_0^{\infty} \frac{1}{1+x^{10}}\,\mathrm{d}x.$

**Solution** The obvious function to use is $f(z) = (1+z^{10})^{-1}$, which is holomorphic except for simple covert poles at the points $\mathrm{e}^{(2k+1)\mathrm{i}\pi/10}$ $(k = 0,\ldots,9)$. We integrate $f$ round the contour $\gamma$ shown in Fig. 20.2 (see the tactical tip below for a discussion of this choice).

Only the pole at $\beta = \mathrm{e}^{\pi\mathrm{i}/10}$ lies inside $\gamma$, and this has residue $1/(10\beta^9) = -\beta/10$ (by 18.7(2)).

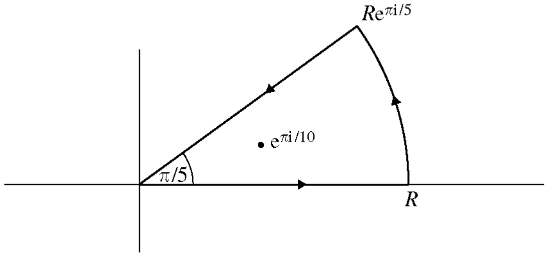



**Figure 20.2** Contour for Example 20.2

On the slanting line, $z = te^{\pi i/5}$ $(0 \leqslant t \leqslant R)$, so $dz/dt = e^{\pi i/5}$ and

$$1 + z^{10} = 1 + t^{10}e^{2\pi i} = 1 + t^{10}.$$

By Cauchy's residue theorem,

$$\int_0^R \frac{1}{1+x^{10}}\,dx + \int_0^{\pi/5} \frac{Rie^{i\theta}}{1+R^{10}e^{10i\theta}}\,d\theta + \int_R^0 \frac{e^{\pi i/5}}{1+t^{10}}\,dt = -2\pi ie^{-\pi i/5}/10.$$

The integral round the circular arc is $\mathbf{O}(R^{-9})$. Letting $R \to \infty$,

$$(1 - e^{\pi i/5})\int_0^\infty \frac{1}{1+x^{10}}\,dx = -\frac{\pi i}{5}e^{\pi i/10}.$$

Now we observe that

$$1 - e^{\pi i/5} = e^{\pi i/10}(e^{-\pi i/10} - e^{\pi i/10}) = -2ie^{\pi i/10}\sin\pi/10.$$

We deduce that

$$\int_0^\infty \frac{1}{1+x^{10}}\,dx = \frac{\pi}{10}\operatorname{cosec}\left(\frac{\pi}{10}\right).$$

**Tactical tips**

- We could have used a semicircular contour. However, had we done so, we would have had to add up the residues at five poles. Our sector contour encloses just one pole.
- Note carefully the way the integral along the slanting line gives a multiple of the integral along the real axis, so that these two integrals combine to give, in the limit, a constant multiple of the integral we want.

When two subpaths in a contour $\gamma$ yield integrals $I$ and $I'$, where $I' = kI$, with $k$ a constant, $k \neq -1$, then we say we have **integral reinforcement**. A simple case of this occurs when $\int_{-R}^{0} f(x)\,\mathrm{d}x = \int_0^R f(x)\,\mathrm{d}x$ (as occurs when $f$ is an even function on $\mathbb{R}$); here the multiplier $k$ is 1. Of course, if $k = -1$ we get cancellation—undesirable if $I$ is the integral we are trying to evaluate!

## Integrals of other functions with a finite number of poles

We consider integrals of the form

$$\int_0^\infty \varphi(x) \left\{ \begin{array}{c} \sin mx \\ \cos mx \end{array} \right\} \mathrm{d}x, \quad \int_{-\infty}^\infty \varphi(x) \left\{ \begin{array}{c} \sin mx \\ \cos mx \end{array} \right\} \mathrm{d}x, \quad \text{and} \quad \int_{-\infty}^\infty \varphi(x)\mathrm{e}^{\pm \mathrm{i}mx}\,\mathrm{d}x,$$

where $m \geqslant 0$ and $\varphi(z) = p(z)/q(z)$ is a rational function where, for now, the polynomials $p$ and $q$ are such that $\deg q > 1 + \deg p$. This restriction ensures that Jordan's inequality is not needed. For examples in which $\deg p < \deg q \leqslant 1 + \deg p$ and Jordan's inequality is used, see Examples 20.14 and 20.16.

**20.3 Example.** To evaluate $\displaystyle\int_{-\infty}^\infty \frac{\cos x}{x^2 + x + 1}\,\mathrm{d}x.$

**Solution** Integrate $f(z) = \mathrm{e}^{\mathrm{i}z}/(z^2 + z + 1)$ round $\gamma = \Gamma(0; R)$, as in Fig. 20.1, with $R > 1$. The real part of $f(z)$, when $z$ is the real variable $x$, is the required integrand. The function $f$ is holomorphic inside and on $\gamma$ except for a simple pole at $z = \omega = \mathrm{e}^{2\pi\mathrm{i}/3}$. By Cauchy's residue theorem and 18.7(2),

$$\int_{-R}^{R} f(x)\,\mathrm{d}x + \int_{\Gamma_R} f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\operatorname{res}\{f(z);\omega\} = 2\pi\mathrm{i}\frac{\mathrm{e}^{\mathrm{i}\omega}}{2\omega + 1} = \frac{2\pi}{\sqrt{3}}\mathrm{e}^{\mathrm{i}(-\frac{1}{2} + \frac{\mathrm{i}}{2}\sqrt{3}\mathrm{i})}.$$

As in 19.7,

$$\left|\int_{\Gamma_R} f(z)\,\mathrm{d}z\right| \leqslant \int_0^\pi \frac{R\mathrm{e}^{-R\sin\theta}}{|R^2\mathrm{e}^{2\mathrm{i}\theta} + R\mathrm{e}^{\mathrm{i}\theta} + 1|}\,\mathrm{d}\theta \leqslant \int_0^\pi \frac{R}{|R^2 - R - 1|}\,\mathrm{d}\theta = \mathbf{O}(R^{-1}).$$

Hence, letting $R \to \infty$ and equating real and imaginary parts in the equation above, we obtain

$$\int_{-\infty}^\infty \frac{\cos x}{x^2 + x + 1}\,\mathrm{d}x = \frac{2\pi}{\sqrt{3}}\mathrm{e}^{-\frac{1}{2}\sqrt{3}} \cos\frac{1}{2}.$$

**Tactical tips**

- ⊙ Observe that $(x^2 + x + 1)^{-1}\cos x$ is not an even function, so that it is not possible to 'double up' to obtain the integral over $[0, \infty)$ rather than that over $(-\infty, \infty)$; contrast this example with that in Example 20.1.

⊙ At first sight it would seem more natural to choose

$$f(z) = \frac{\cos z}{z^2 + z + 1}$$

rather than introducing the complex exponential $e^{iz}$ as we did. However this does not work. Our method relies on our showing that $\int_{\Gamma_R} f(z)\,dz \to 0$ as $R \to \infty$. On $\Gamma_R$ we have

$$\left|e^{iRe^{i\theta}}\right| = e^{-R\sin\theta} \leqslant 1,$$

whereas, by Exercise 7.6,

$$\left|\cos(Re^{i\theta})\right|^2 = \cosh^2(R\sin\theta) - \sin^2(R\cos\theta),$$

and this grows like $\cosh^2 R$ when $\theta$ is close to $\pi/2$. Thus if we had chosen $f(z) = (z^2+z+1)^{-1}\cos z$ we could not have shown that $\int_{\Gamma_R} f(z)\,dz \to 0$ as $R \to \infty$.

⊙ Note that we saw that the integral round the large arc is $\mathbf{O}(R^{-1})$ using just the inequality $e^{-R\sin\theta} \leqslant 1$ for $\theta \in [0,\pi]$. Jordan's inequality (19.5) is *not* needed here.

**20.4 Example.** To evaluate $\displaystyle\int_0^\infty \frac{\sin^2 x}{x^2}\,dx.$

**Solution** For the reasons explained in the preceding tactical tip, we cannot take $f(z) = \sin^2 z/z^2$. We would like instead a function involving a complex exponential whose real part is $\sin^2 x/x^2$ when $z = x$ is real. Recall that $2\sin^2 x = 1 - \cos 2x$. This suggests taking $f(z) = (1 - e^{2iz})/z^2$. The function $f$ is holomorphic except for a pole at $0$. The Laurent expansion about $0$ is

$$\frac{1 - e^{2iz}}{z^2} = \frac{1}{z^2}\left(1 - (1 + 2iz + \dots)\right),$$

so that the pole is simple, with residue $-2i$. We cannot use the semicircular contour $\Gamma(0;R)$ because it passes through a pole. Instead, we take the contour $\gamma$ shown in Fig. 20.3; $f$ is holomorphic inside and on $\gamma$, so, by Cauchy's theorem,

$$\int_{-R}^{-\varepsilon} f(x)\,dx - \int_{\Gamma_\varepsilon} f(z)\,dz + \int_{\varepsilon}^{R} f(x)\,dx + \int_{\Gamma_R} f(z)\,dz = 0.$$

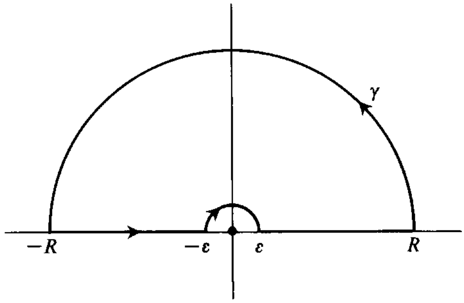



**Figure 20.3** Contour for Example 20.4

The first and third integrals combine to give

$$\int_{\varepsilon}^{R} \frac{1-\mathrm{e}^{-2\mathrm{i}x}}{x^2}\,\mathrm{d}x + \int_{\varepsilon}^{R} \frac{1-\mathrm{e}^{2\mathrm{i}x}}{x^2}\,\mathrm{d}x = 2\int_{\varepsilon}^{R} \frac{1-\cos 2x}{x^2}\,\mathrm{d}x = \int_{\varepsilon}^{R} \frac{4\sin^2 x}{x^2}\,\mathrm{d}x.$$

Because the pole at 0 is simple, the Indentation lemma (19.9) is applicable. It implies that

$$\lim_{\varepsilon\to 0}\int_{\Gamma_\varepsilon} f(z)\,\mathrm{d}z = \mathrm{i}(\pi-0)\mathrm{res}\,\{f(z);0\} = 2\pi.$$

Also

$$\left|\int_{\Gamma_R} f(z)\,\mathrm{d}z\right| \leqslant \int_0^{\pi} \frac{1+\mathrm{e}^{-2R\sin\theta}}{R^2} R\,\mathrm{d}\theta = \mathbf{O}(R^{-1}).$$

Letting $R\to\infty$ and $\varepsilon\to 0$,

$$\int_0^{\infty} \frac{\sin^2 x}{x^2}\,\mathrm{d}x = \frac{\pi}{2}.$$

**Tactical tip**

⊙ Care is needed with the signs when handling the indentation integral. Remember that the small arc is traced clockwise (hence the minus sign preceding the integral along $\Gamma_\varepsilon$). Note, too, that the contribution from the residue at 0 comes from the integral round the contour; it is not a term in a sum of residues. It is instructive to re-work the last example using an indented semicircle with the indentation in the lower half-plane and enclosing the pole. You should get the same answer!

**20.5 Example (semicircles: up versus down).** To prove that

$$\int_0^\infty \frac{e^{-2ix}}{1+x^4}\,dx = 2\sqrt{2}\pi(\cos\sqrt{2}-\sin\sqrt{2})e^{-1/\sqrt{2}}.$$

**Solution** The natural choice of function is $f(z) = e^{-2iz}(1+z^4)^{-1}$, which is holomorphic except for simple poles at $e^{(2k+1)\pi i}$ $(k = 0, 1, 2, 3)$. When $z = Re^{i\theta}$, we have $\left|e^{-2iz}\right| = e^{2R\sin\theta}$, and this is not bounded in the upper half-plane, where $\sin\theta \geqslant 0$. We therefore cannot use the semicircle $\Gamma(0; R)$ because the integral of $f(z)$ round $\Gamma_R$ does not tend to zero. However,

$$\left|e^{-2iz}\right| = e^{2R\sin\theta} \leqslant 1 \quad \text{provided } -\pi \leqslant \theta \leqslant 0.$$

So we take our contour $\gamma$ to be a semicircle in the *lower* half-plane.

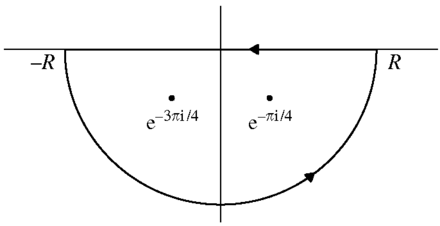



**Figure 20.4** Example 20.5: go down not up

By Cauchy's residue theorem we have (note that the real axis is traced from right to left),

$$\begin{aligned}
-\int_{-R}^{R} \frac{e^{-2ix}}{1+x^4}\,dx + \int_{-\pi}^{0} \frac{e^{-2iRe^{i\theta}}}{1+R^4e^{4i\theta}} Rie^{i\theta}\,d\theta
&= 2\pi i\left(\operatorname{res}\left\{f(z); e^{-i\pi/4}\right\} + \operatorname{res}\left\{f(z); e^{-3i\pi/4}\right\}\right)\\
&= 2\pi i\left[\frac{-z^3}{4}e^{-2iz}\right]_{z=e^{-i\pi/4}=\frac{1}{\sqrt{2}}(1-i)}\\
&\quad + 2\pi i\left[\frac{-z^3}{4}e^{-2iz}\right]_{z=e^{-3i\pi/4}=\frac{1}{\sqrt{2}}(-1-i)}.
\end{aligned}$$

The integral round the circular arc is $\mathbf{O}(R^{-3})$ because the exponential factor is bounded by 1. Hence, simplifying the sum of residues and letting $R \to \infty$, we obtain the result required.

**Tactical tip**

⊙ If we have a factor of $e^{\pm imx}$ with an unspecified real value of $m$, it is necessary to treat the cases $m \geqslant 0$ and $m < 0$ separately, taking different contours in the two cases. See 22.11 for an example of this sort.

## Integrals involving functions with infinitely many poles

The method of this section can be used to evaluate integrals of the type

$$\int_{-\infty}^{\infty} \varphi(x) \left\{ \begin{matrix} \sin mx \\ \cos mx \end{matrix} \right\} dx,$$

where $\varphi(z)$ is a function which has an infinite number of regularly spaced poles. Examples of such functions are $\operatorname{cosec} z$, $\operatorname{sech} z$, and $(1-e^z)^{-1}$.

**20.6 Example.** To evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{e^{ax}}{\cosh x}\, dx \quad (-1 < a < 1)$.

(The restriction on $a$ is needed for the integral to exist.)

**Solution** The function $f(z) = e^{az}\operatorname{sech} z$ has simple poles at $z = \frac{1}{2}(2n+1)\pi i$ $(n \in \mathbb{Z})$. It is holomorphic inside and on the contour $\gamma$ shown in Fig. 20.5 except at $\pi i/2$, inside $\gamma$, where there is a simple pole of residue $-ie^{\frac{1}{2}a\pi i}$ (by 18.7(2)). By Cauchy's residue theorem,

$$\int_{-S}^{R} \frac{e^{ax}}{\cosh x}\, dx + \int_{0}^{\pi} \frac{e^{a(R+iy)}}{\cosh(R+iy)}\, i\, dy + \int_{R}^{-S} \frac{e^{a\pi i}e^{ax}}{\cosh(x+\pi i)}\, dx$$
$$+ \int_{\pi}^{0} \frac{e^{a(-S+iy)}}{\cosh(-S+iy)}\, dy = 2\pi e^{a\pi i}.$$

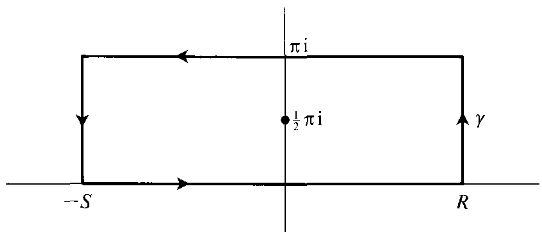



**Figure 20.5** Contour for Example 20.6

Consider the integrals along the vertical lines:

$$\left|\int_0^\pi \frac{e^{a(R+iy)}}{\cosh(R+iy)}\,i\,dy\right| \leqslant \int_0^\pi \left|\frac{2e^{a(R+iy)}}{e^{(R+iy)}+e^{-(R+iy)}}\right| dy$$
$$\leqslant \int_0^\pi \frac{2e^{aR}}{|e^R - e^{-R}|}\,dy \to 0 \text{ as } R \to \infty \qquad (\text{since } a < 1)$$

and

$$\left|\int_0^\pi \frac{e^{a(-S+iy)}}{\cosh(-S+iy)}\,i\,dy\right| \leqslant \int_0^\pi \left|\frac{2e^{a(-S+iy)}}{e^{(-S+iy)}+e^{-(-S+iy)}}\right| dy$$
$$\leqslant \int_0^\pi \frac{2e^{-aS}}{|e^{-S} - e^{S}|}\,dy \to 0 \text{ as } S \to \infty \qquad (\text{since } a > -1).$$

Letting $R$ and $S$ tend to infinity, we obtain

$$\int_{-\infty}^{\infty} \frac{e^{ax}}{\cosh x}\,dx = \frac{2\pi e^{\frac{1}{2}a\pi i}}{1+e^{a\pi i}} = \frac{2\pi}{e^{-\frac{1}{2}a\pi i}+e^{\frac{1}{2}a\pi i}} = \pi \sec\left(\frac{\pi a}{2}\right).$$

**Tactical tips** The following observations motivate our choice of contour.

- ⊙ A semicircle would have involved, in the limit, an infinite sum of residues.
- ⊙ The horizontal sides of the contour avoid the poles. One is along $y = 0$, to give the integral we want; the other is along $y = \pi$ and yields a multiple of the required integral because $\cosh(x + \pi i) = -\cosh x$ and $e^{a(x+\pi i)} = e^{a\pi i}e^{ax}$. This is another instance of integral reinforcement (recall the tactical tip in 20.2).
- [For the benefit of those who care about the sense in which integrals exist: taking the vertical sides of the rectangle along $x = -S, R$, rather than along $x = \pm R$, leads directly to the improper integral over $(-\infty, \infty)$ rather than to a principal-value integral (see 19.15).]

## Integrals involving multifunctions

We consider integrals of the form

$$\int_0^\infty \varphi(x)\log x\,\mathrm{d}x \quad \text{and} \quad \int_0^\infty \varphi(x)x^{a-1}\,\mathrm{d}x \quad (a>0),$$

where $\varphi(z)$ is meromorphic. Since logarithms and non-integer powers are multifunctions once we move into the complex plane, we work in a cut plane, with a selected holomorphic branch of the multifunction. The branch point at 0 is avoided by means of an indentation. Recall the discussion of integration in cut planes given in 19.11; this ensures that we can legitimately use Cauchy's theorem and Cauchy's residue theorem, even when our contour goes along an edge of the cut. We give two examples here; for others see 21.12 and 22.13.

**20.7 Example.** To evaluate $\displaystyle\int_0^\infty \frac{\log x}{1+x^2}\,\mathrm{d}x$.

**Solution** We cut the plane along $(-\infty, 0]$ and take the holomorphic branch of the logarithm given by $\log z = \log|z| + \mathrm{i}\theta$, where $z = |z|\mathrm{e}^{\mathrm{i}\theta}$ and $\theta$ is between $-\pi$ and $\pi$. Then $f(z) = (1+z^2)^{-1}\log z$ is holomorphic in the cut plane except for simple poles at $\pm\mathrm{i}$. Let $\gamma$ be the contour in Fig. 20.6. On the top side of the cut, $\theta = \pi$, so $\log z = \log x + \mathrm{i}\pi$, where $-z = x > 0$. Cauchy's residue theorem gives

$$\int_\varepsilon^R \frac{\log x}{1+x^2}\,\mathrm{d}x + \int_{\Gamma_R} f(z)\,\mathrm{d}z + \int_R^\varepsilon \frac{\log x + \mathrm{i}\pi}{1+x^2}(-\,\mathrm{d}x) - \int_{\Gamma_\varepsilon} f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\,\mathrm{res}\,\{f(z);\mathrm{i}\}.$$

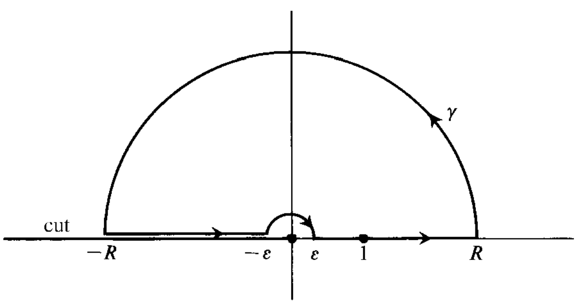



**Figure 20.6** Contour for Example 20.7

We have

$$\operatorname{res}\{f(z); \mathrm{i}\} = \frac{\log \mathrm{i}}{2\mathrm{i}} = \frac{\frac{1}{2}\pi \mathrm{i}}{2\mathrm{i}}.$$

Also,

$$\begin{aligned}\left|\int_{\Gamma_R} f(z)\,\mathrm{d}z\right| &\leqslant \int_0^{\pi} \left|\frac{\log R + \mathrm{i}\theta}{1 + R^2 \mathrm{e}^{2\mathrm{i}\theta}} R\mathrm{i}\mathrm{e}^{\mathrm{i}\theta}\right| \mathrm{d}\theta \\ &\leqslant \int_0^{\pi} \frac{(\log R + \pi)R}{R^2 - 1}\,\mathrm{d}\theta = \mathbf{O}(R^{-1}\log R)\end{aligned}$$

and

$$\left|\int_{\Gamma_\varepsilon} f(z)\,\mathrm{d}z\right| \leqslant \int_0^{\pi} \frac{(|\log \varepsilon| + \pi)\varepsilon}{1 - \varepsilon^2}\,\mathrm{d}\theta = \mathbf{O}(\varepsilon \log \varepsilon).$$

Invoking the basic limits in 19.6, we get, as $R \to \infty$ and $\varepsilon \to 0$,

$$2\int_0^{\infty} \frac{\log x}{1 + x^2}\,\mathrm{d}x + \mathrm{i}\pi \int_0^{\infty} \frac{1}{1 + x^2}\,\mathrm{d}x = \frac{1}{2}\pi^2 \mathrm{i}.$$

By equating real parts we get

$$\int_0^{\infty} \frac{\log x}{1 + x^2}\,\mathrm{d}x = 0.$$

(Equating imaginary parts gives us an integral we do not need contour integration to compute!)

**Tactical tip**

- Note how in the example above the two integrals along subintervals of the real axis reinforce rather than cancelling. Compare this with the situation in the next example, where integrals taken along the two edges of a cut reinforce. The need for reinforcement, rather than cancellation, often governs the choice of contour in multifunction examples such as these.

**20.8 Example.** To prove that $\displaystyle\int_0^{\infty} (1 + x^3)^{-1}\sqrt{x}\,\mathrm{d}x = \frac{\pi}{3}.$

**Solution** The function $(1 + z^3)^{-1}$ has simple poles at $-1$, $\mathrm{e}^{\pi \mathrm{i}/3}$, and $\mathrm{e}^{-\pi \mathrm{i}/3}$. Because there is a pole at $-1$, we control the square root by cutting the plane along $[0, \infty)$ rather than along $(-\infty, 0]$. We take the holomorphic branch $z^{1/2} = |z|^{1/2}\,\mathrm{e}^{\mathrm{i}\theta/2}$, with $\theta$ between 0 and $2\pi$, and use the contour shown in Fig. 20.7. On the top side of the cut, $z = x > 0$ and $z^{1/2} = x^{1/2}$, while on the bottom side, $z = |z|\,\mathrm{e}^{2\pi \mathrm{i}}$ and $z^{1/2} = -x^{1/2}$, where $x > 0$. Since the integrals along the two

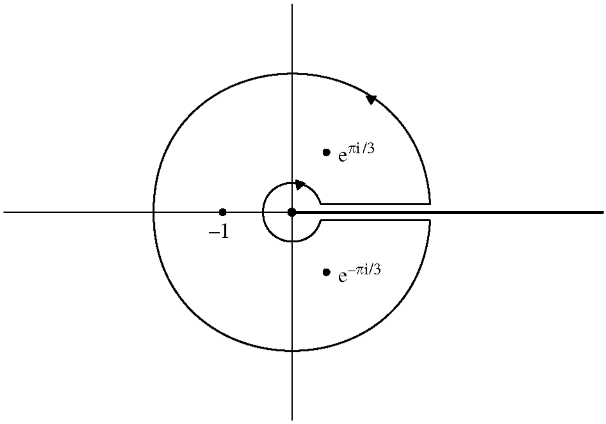



**Figure 20.7** Contour for Example 20.8

sides of the cut are in opposite directions, the integrals reinforce. The remainder of the calculation involves no new techniques and we leave it as an exercise.

## Evaluation of definite integrals: overview (basic track)

**20.9 Making choices.** Suppose you are confronted by a real definite integral $I$. How do you decide how to choose a function $f$ and a contour $\gamma$ so that evaluating $\int_\gamma f(z)\,dz$ leads you to the value of $I$? There are some guiding principles:

- $\int_\gamma f(z)\,dz$, or its real or imaginary part, must incorporate $I$ or a quantity converging to $I$.
- Cauchy's theorem or Cauchy's residue theorem must be applicable, so $f$ has to be holomorphic except for poles, none on the contour and only finitely many inside it.
- Indentations may be used to avoid simple poles, but you cannot indent to avoid a *multiple* pole.
- When dealing with logarithms and non-integer powers, the plane should be cut appropriately and a selected branch of a multifunction chosen, the contour must not cross the cut(s), and indentations must be used to avoid any branch point(s). Note that a branch must be specified by arg restrictions; these restrictions need to match the position of the cut(s); recall 9.9.

It is usually clear what interval(s) of the real axis to include in the contour. When deciding how to join up the endpoints to form a suitable contour, remember that the integral along each added path is normally handled in one of the following ways:

- estimation, to give zero limiting value or a limiting value we can compute;
- integral reinforcement.

**20.10 Complex substitutions are not allowed.** Consider the following erroneous argument. Let $I = \int_0^\infty (1+x^4)^{-1}\,\mathrm{d}x$. Put $x = \mathrm{i}y$. Then

$$I = \int_0^\infty (1+y^4)^{-1}\,\mathrm{d}y = \mathrm{i}I.$$

So $I = 0$, which is clearly wrong, because the integrand is strictly positive. The correct value for $I$ is $\pi/(2\sqrt{2})$, as we showed in Example 19.1 by integrating $(1+z^4)^{-1}$ round a semicircular contour. We could alternatively have derived this result using the contour in Fig. 20.8. We can now see that the integrals along the rays $\arg z = 0$ and $\arg z = \pi/2$, which the substitution equates, in fact differ by $2\pi\mathrm{i}\,\mathrm{res}\{(1+z^4)^{-1}; \mathrm{e}^{\pi\mathrm{i}/4}\}$. To assert this, we have to prove also that the integral along the linking arc tends to zero as $R \to \infty$. For examples of a similar kind, see 20.16 and 22.12.

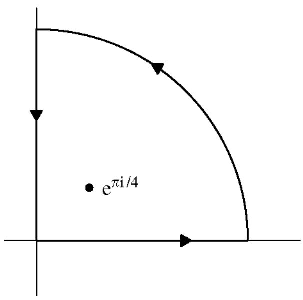



**Figure 20.8** The dangers of complex substitution

In conclusion: by making a complex substitution you may overlook, at your peril, residues and/or necessary linking arcs. Don't be tempted by complex substitution. Apply Cauchy's theorem or the residue theorem instead!

## Summation of series

Convergence tests establish that certain series converge to finite limits, but they do not yield the value of the sum. If an infinite sum can be recognized as a sum of residues of a meromorphic function, then contour integration may enable us to evaluate it. Before discussing the general method we present a typical example.

**20.11 Example.** To prove that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

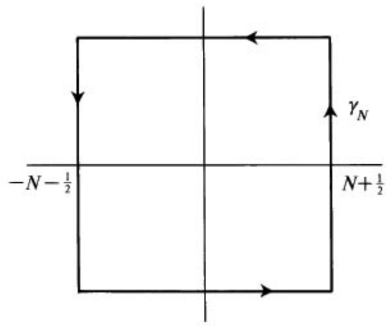



**Figure 20.9** Contour for Example 20.11

**Solution** The function $f(z) = \pi z^{-2} \cot \pi z$ is holomorphic except for simple covert poles at $n$ $(n = \pm 1, \pm 2, \ldots)$ of residue $1/n^2$ and a triple pole at 0 of residue $-\pi^2/3$ (see 18.10). Integrate $f(z)$ round the square contour $\gamma_N$ shown in Fig. 20.9; $\gamma_N^*$ is the square $S_N$ with vertices at $(\pm 1 \pm \mathrm{i})(N + \frac{1}{2})$. Note that $f$ is holomorphic inside and on $\gamma_N$ except for poles at $0, \pm 1, \ldots, \pm N$. By Cauchy's residue theorem,

$$\int_{\gamma_N} f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\left(2\sum_{n=1}^{N} \frac{1}{n^2} - \frac{\pi^2}{3}\right).$$

It is now enough to show that $\int_{\gamma_N} f(z)\,\mathrm{d}z \to 0$ as $N \to \infty$. We have

$$\begin{aligned}\left|\int_{\gamma_N} f(z)\,\mathrm{d}z\right| &\leqslant \sup_{z \in S_N} \left|\frac{\pi \cot \pi z}{z^2}\right| \times \operatorname{length}(\gamma_N) \\ &\leqslant \sup_{z \in S_N} |\cot \pi z| \,\frac{4(2N+1)\pi}{(N+\frac{1}{2})^2}.\end{aligned}$$

This is $\mathbf{O}(N^{-1})$, by the Boundedness lemma for cot, 19.13(1).

**20.12 The summation of series by contour integration.** The method used in Example 20.11 applies to any sum $\sum_{n=1}^{\infty} \phi(n)$, where $\phi$ is a function with the following properties:

(i) $\phi(n) = \phi(-n)$ for all $n = 1, 2, \dots$,
(ii) $\phi$ is a rational function, and
(iii) $\phi(z) = \mathbf{O}(|z|)^{-2}$ for large $|z|$.

We integrate $f(z) = \phi(z)\pi \cot \pi z$ round the contour $\gamma_N$ in Fig. 20.9 and use Cauchy's residue theorem. The term $\pi \cot \pi z$ creates simple poles of $f$ at each $n \in \mathbb{Z}$ at which $\phi$ is holomorphic and non-zero, of residue $\phi(n)$. The bound on $\phi$, combined with the Boundedness lemma for cot (19.13(1)), ensures that $\int_{\gamma_N} f(z)\, \mathrm{d}z \to 0$ as $N \to \infty$.

Under the same conditions (i)–(iii) on $\phi$, we can evaluate $\sum_{n=1}^{\infty}(-1)^n \phi(n)$ by integrating $f(z) = \phi(z)\pi \operatorname{cosec} \pi z$ round the same square contours as before. The cosec term creates a simple covert pole of residue $(-1)^n\phi(n)$ at $n$ (provided $\phi$ is holomorphic and non-zero there). We then invoke Cauchy's residue theorem and the Boundedness lemma for cosec (19.13(2)).

An adaptation of these techniques can be used to obtain series expansions of certain meromorphic functions. See Exercise 20.11 for a typical example.

## Further techniques

This optional section presents more subtle and more ingenious techniques which enlarge the range of integrals which can be evaluated by contour integration. We shall comment only on new features, leaving routine verification of residues and estimates as exercises.

**20.13 Example (avoidance of a multiple pole at an indentation).** To prove

$$\int_0^{\infty} \frac{x - \sin x}{x^3}\, \mathrm{d}x = \frac{\pi}{4}.$$

**Solution** The integrand is the real part, when $z = x$ is real, of $z^{-3}(z + \mathrm{i}e^{\mathrm{i}z})$. There is certainly a pole at 0, so we would hope to integrate round an indented semicircle. However the Laurent expansion

$$z^{-3}(z + \mathrm{i}e^{\mathrm{i}z}) = z^{-3}\left(z + \mathrm{i}(1 + \mathrm{i}z + \tfrac{1}{2}(\mathrm{i}z)^2) + \dots\right)$$

reveals a triple pole. This prevents us applying the Indentation lemma. It is the term $\mathrm{i}z^{-3}$ in the expansion that causes the pole to be multiple. We eliminate it by taking $f(z) = z^{-3}(z + \mathrm{i}e^{\mathrm{i}z} - \mathrm{i})$, which still gives the required integrand when $z$ is real. The calculation now proceeds as in 20.4.

**20.14 Example (use of Jordan's inequality).** To prove that $\displaystyle\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}$.

[The integral exists only as an improper integral.]

**Solution** We integrate $f(z) = \mathrm{e}^{\mathrm{i}z}/z$ round the indented semicircle shown in Fig. 20.3. (Remember the tactical tip in 20.3, which explains why we use this function and not $(\sin z)/z$.) We now proceed as in Example 20.14, except for one important difference. Putting the usual upper bound $\sin\theta \leqslant 1$ into the estimate

$$\left|\int_{\Gamma_R} f(z)\,\mathrm{d}z\right| \leqslant \int_0^\pi \mathrm{e}^{-R\sin\theta}\,\mathrm{d}\theta$$

does not show that the integral on the right-hand side tends to zero. We need the tighter estimate supplied by Jordan's inequality. As in 19.7 we obtain

$$\left|\int_{\Gamma_R} f(z)\,\mathrm{d}z\right| \leqslant 2\int_0^{\pi/2} \mathrm{e}^{-R\sin\theta}\,\mathrm{d}\theta \leqslant \int_0^{\pi/2} \mathrm{e}^{-2R\theta/\pi}\,\mathrm{d}\theta = \mathbf{O}(R^{-1}).$$

The pole of $f$ at 0 is simple. The Indentation lemma gives the limit of the integral round $-\Gamma_\varepsilon$: the limit is $\mathrm{i}(\pi - 0)\mathrm{res}\{f(z); 0\} = \mathrm{i}\pi$. Finally we note that the integrals of $f$ along the positive and negative real axes combine to give $2\mathrm{i}\int_0^\infty (\sin x)/x\,\mathrm{d}x$.

**20.15 Example (a ubiquitous integral).** To prove that $\displaystyle\int_0^\infty \mathrm{e}^{-x^2}\,\mathrm{d}x = \sqrt{\pi}/2$.

(This well-known fact is usually obtained by methods other than contour integration. To add variety to our catalogue of contour integrals we show how, with some ingenuity, the result can be obtained from the residue theorem.)

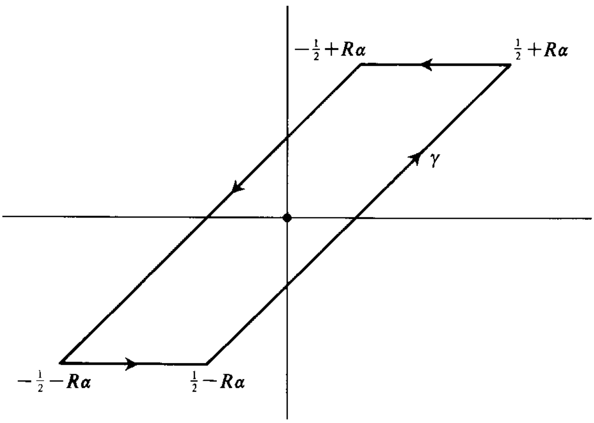



**Figure 20.10** Contour for Example 20.15

**Solution** Let $\alpha = \mathrm{e}^{\mathrm{i}\pi/4}$. We integrate $f(z) = \mathrm{e}^{\mathrm{i}\pi z^2} \operatorname{cosec} \pi z$ round the contour $\gamma$ in Fig. 20.10 and obtain

$$\int_\gamma f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\,\mathrm{res}\,\{f(z); 0\} = 2\mathrm{i}.$$

On the slanting sides, we have $z = t\alpha \pm \frac{1}{2}$ $(-R \leqslant t \leqslant R)$ and $f(z) = \pm(\sec \pi t\alpha)\mathrm{e}^{\mathrm{i}\pi(t^2 \pm t\alpha + \frac{1}{4})}$, so that their combined contribution to $\int_\gamma f(z)\,\mathrm{d}z$ is

$$\int_{-R}^{R} \alpha(\mathrm{e}^{\mathrm{i}\pi(t^2+t\alpha+\frac{1}{4})} + \mathrm{e}^{\mathrm{i}\pi(t^2-t\alpha+\frac{1}{4})}) \sec \pi t\alpha\,\mathrm{d}t$$

$$= 2\mathrm{i}\int_{-R}^{R} \mathrm{e}^{-t^2}\,\mathrm{d}t = \frac{4\mathrm{i}}{\sqrt{\pi}}\int_0^{R\sqrt{\pi}} \mathrm{e}^{-x^2}\,\mathrm{d}x.$$

On the horizontal sides, $z = \pm R\alpha + t$ $(-1/2 \leqslant t \leqslant 1/2)$ and an estimate of their contribution to the integral is given by

$$\left|\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\mathrm{e}^{\mathrm{i}\pi(\mathrm{i}\alpha^2 \pm 2R\alpha t + t^2)}}{\sin \pi(\pm R\alpha + t)}\,\mathrm{d}t\right| \leqslant \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{2\mathrm{e}^{-\pi R^2 + R\pi t\sqrt{2}}}{\mathrm{e}^{\pi R/\sqrt{2}} - \mathrm{e}^{-\pi R/\sqrt{2}}}\,\mathrm{d}t.$$

Taking the limit as $R \to \infty$, the required result is obtained.

We have already indicated that contour integration can sometimes provide a substitute for making a substitution. Our final example illustrates how this technique allows us to derive other integrals from the known integral $\int_0^\infty \mathrm{e}^{-x^2}\,\mathrm{d}x$.

**20.16 Example (deduction from known integral; Jordan's inequality).** To prove that $\displaystyle\int_0^\infty \cos x^2\,\mathrm{d}x = \int_0^\infty \sin x^2\,\mathrm{d}x = \sqrt{\pi/8}$. [This exists only as an improper integral.]

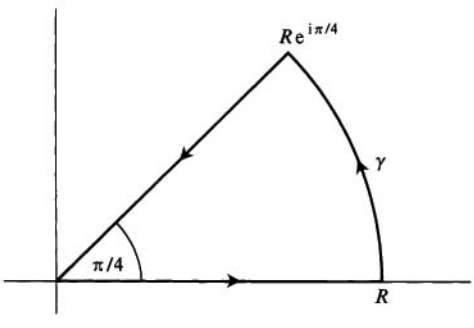



**Figure 20.11** Contour for Example 20.16

**Solution** Integrate $f(z) = e^{iz^2}$ round the contour $\gamma$ shown in Fig. 20.11. By Cauchy's theorem,

$$\int_0^R e^{ix^2}\,dx + \int_0^{\pi/4} e^{iR^2e^{2i\theta}} Rie^{i\theta}\,d\theta + \int_R^0 e^{i(te^{i/4})^2} e^{\pi i/4}\,d\theta = 0.$$

By Jordan's inequality (19.5), $\sin 2\theta \geqslant 4\theta/\pi$ for $\theta \in [0, \pi/4]$, so

$$\left|\int_0^{\pi/4} e^{iR^2e^{2i\theta}} Rie^{i\theta}\,d\theta\right| \leqslant R\int_0^{\pi/4} Re^{-R^2\sin 2\theta}\,d\theta$$
$$\leqslant R\int_0^{\pi/4} Re^{-4R^2\theta/\pi}\,d\theta = \frac{\pi}{4R}(1 - e^{-R^2}).$$

Hence, letting $R \to \infty$,

$$\int_0^\infty (\cos x^2 + i\sin x^2)\,dx = \int_0^\infty e^{ix^2}\,dx = \frac{1+i}{\sqrt{2}}\int_0^\infty e^{-t^2}\,dt = \frac{(1+i)\sqrt{\pi}}{2\sqrt{2}}.$$

Now equate real and imaginary parts.

**Tactical tip**

- ⊙ If we had, naively, made the complex substitution of $\alpha x$ for $x$ (with $\alpha = e^{i\pi/4}$) in $\int_0^\infty e^{-x^2}\,dx$, we would, ostensibly, have obtained the desired result. The substitution equates the integrals along the two rays $\arg z = 0$ and $\arg z = \pi/4$. Looking at the contour integral, we see that these integrals are in fact equal only because there are no poles of $e^{iz^2}$ in the sector between the rays and the integral along the circular arc tends to zero as $R \to \infty$.

## Exercises

**Exercises from the text.** Complete the calculations in 20.8. Using a right-angled sector contour, evaluate $\int_0^\infty (x^4+1)^{-1}\,dx$ (see 20.10)).

20.1 Prove that

(i) $$\int_0^\infty \frac{1}{(x^2+a^2)(x^2+b^2)}\,dx = \frac{\pi}{2ab(a+b)} \quad (a, b > 0,\ a \neq b),$$

(ii) $$\int_{-\infty}^\infty \frac{1}{(x^2+x+1)^2}\,dx = \frac{4\pi}{3\sqrt{3}},$$

(iii) $$\int_{-\infty}^\infty \frac{1}{x^4+x^2+1}\,dx = \frac{\pi}{\sqrt{3}}.$$

20.2 (a) By integrating $(1+z^n)^{-1}$ round a suitable sector of angle $2\pi/n$, prove that, for $n = 2, 3, \ldots,$

$$\int_0^\infty (1+z^n)^{-1}\,\mathrm{d}x = \frac{\pi}{n}\operatorname{cosec}\left(\frac{\pi}{n}\right).$$

(b) Evaluate also $\int_0^\infty x(1+x^n)^{-1}\,\mathrm{d}x$ $(n = 2, 3, \ldots)$.

(c) Could a semicircular contour be used in either (a) or (b)?

20.3 Prove that

$$\text{(i)}\ \int_{-\infty}^\infty \frac{\cos x}{x^2+a^2}\,\mathrm{d}x = \frac{\pi}{a}\mathrm{e}^{-a} \quad (a>0), \qquad \text{(ii)}\ \int_0^\infty \frac{\cos x}{(x^2+1)^2}\,\mathrm{d}x = \frac{\pi}{\mathrm{e}},$$

$$\text{(iii)}\ \int_{-\infty}^\infty \frac{\cos \pi x}{x^2-2x+2}\,\mathrm{d}x = -\pi\mathrm{e}^{-\pi}.$$

(Note: Jordan's inequality is not needed in this exercise or the next one.)

20.4 Prove that

$$\text{(i)}\ \int_0^\infty \frac{\cos ax - \cos bx}{x^2}\,\mathrm{d}x = \frac{\pi}{2}(b-a) \quad (a, b > 0),$$

$$\text{(ii)}\ \int_0^\infty \frac{\sin x}{x(x^2+1)}\,\mathrm{d}x = \frac{\pi}{2}(1+\cos 1).$$

20.5 Evaluate $\int_{-\infty}^\infty \frac{\mathrm{e}^{ax}}{1+\mathrm{e}^x}\,\mathrm{d}x$ and deduce the value of $\int_0^\infty \frac{x^{a-1}}{1+x}\,\mathrm{d}x$.

20.6 By integrating a suitably chosen branch of $z^{a-1}/(1+z)$ round the contour in Fig. 20.7, prove that

$$\int_0^\infty \frac{x^{a-1}}{1+x}\,\mathrm{d}x = a\operatorname{cosec}\pi a \quad (0 < a < 1).$$

20.7 Prove that

$$\text{(i)}\ \int_0^\infty \frac{\log x}{1+x^4}\,\mathrm{d}x = \frac{\pi^2}{8\sqrt{2}}, \qquad \text{(ii)}\ \int_0^\infty \frac{(\log x)^2}{1+x^2}\,\mathrm{d}x = \frac{\pi^3}{8}.$$

20.8 Prove that

$$\text{(i)}\ \sum_{n=-\infty}^{\infty} \frac{1}{n^2+1} = \pi\coth\pi,$$

$$\text{(ii)}\ \sum_{n=-\infty}^{\infty} \frac{1}{(n-a)^2} = \pi^2\operatorname{cosec}^2\pi a \quad (a \notin \mathbb{Z}).$$

20.9 By integrating suitable functions $\phi(z)\operatorname{cosec}\pi z$ (see 20.12), prove that

$$\text{(i)} \sum_{n=-\infty}^{\infty} (-1)^n(n^2+1)^{-1} = \pi \operatorname{cosech}\pi, \qquad \text{(ii)} \sum_{n=1}^{\infty}(-1)^n n^{-2} = -\frac{\pi^2}{12},$$

20.10 Let the Taylor expansion of $\pi z \cot \pi z$ about 0 be $\sum_{n=0}^{\infty} c_n z^n$. Prove that

$$c_{2n} = -2\sum_{k=1}^{\infty} k^{-2n} \quad (n \geqslant 1).$$

20.11 By integrating $f(w) = \dfrac{\operatorname{cosec} w}{w(w-z)}$ round a suitable contour, prove that

$$\operatorname{cosec} z = \frac{1}{z} - 2z\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2\pi^2 - z^2} \quad (z \neq k\pi \ (k \in \mathbb{Z})).$$

20.12 Use a method similar to that in the preceding exercise to obtain the following expansions:

$$\text{(i)} \ \pi \cot \pi z = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{z}{z^2 - n^2} \quad (z \neq k \ (k \in \mathbb{Z})),$$

$$\text{(ii)} \ \frac{1}{e^z - 1} = \frac{1}{z} - \frac{1}{2} + \sum_{n=1}^{\infty} \frac{2z}{z^2 + 4n^2\pi^2} \quad (z \neq 2m\pi i \ (m \in \mathbb{Z})).$$

20.13 (A miscellany of integrals, to test ability to identify and apply the right techniques (basic track only).) Prove that

$$\text{(i)} \int_0^{\infty} \frac{x^2}{1+x^4}\,dx = \frac{\pi}{2\sqrt{2}},$$

$$\text{(ii)} \int_0^{\infty} \frac{\sin x}{x^2+x+1}\,dx = -\frac{2\pi}{\sqrt{3}} \sin \tfrac{1}{2} e^{-\frac{\sqrt{3}}{2}},$$

$$\text{(iii)} \int_0^{\infty} \frac{\log x}{(1+x^2)^2}\,dx = -\frac{\pi}{4},$$

$$\text{(iv)} \int_0^{\infty} \frac{\cosh ax}{\cosh x}\,dx = \frac{\pi}{2}\sec\left(\frac{\pi a}{2}\right) \quad (-1 < a < 1).$$

20.14 Consider evaluating the following integrals by applying Cauchy's theorem or Cauchy's residue theorem to $\int_\gamma f(z)\,dz$ for suitably chosen $f$ and $\gamma$.

(a) State what function $f$ and contour $\gamma$ you would choose;

(b) indicate the location and type of any singularities of $f$;

(c) where appropriate, comment on how the tactic of integral reinforcement would be employed in arriving at the desired integral;

(d) comment on the estimation of any integrals round large or small arcs.

$$\text{(i)} \int_0^\infty \frac{x^2}{(x^4+a^4)^3}\,\mathrm{d}x \quad (a \in \mathbb{R}), \qquad \text{(ii)} \int_0^\infty \frac{x^2}{1+x^{12}}\,\mathrm{d}x,$$

$$\text{(iii)} \int_0^\infty \frac{\log x}{(x-a)^2+b^2}\,\mathrm{d}x \quad (a, b > 0), \qquad \text{(iv)} \int_0^\infty \frac{1}{(1+x^2)^n}\,\mathrm{d}x,$$

$$\text{(v)} \int_0^\infty \frac{\mathrm{e}^{-imx}}{(1+x^2)^4}\,\mathrm{d}x \quad (m \in \mathbb{R}), \qquad \text{(vi)} \int_0^\infty \frac{1-\cos x}{x^2(1+x^4)}\,\mathrm{d}x,$$

$$\text{(vii)} \int_0^\infty \frac{x^a}{(1+x)^2}\,\mathrm{d}x \quad (-1 < a < 1), \qquad \text{(viii)} \int_{-\infty}^\infty \frac{\cos(\pi x/2)}{x^2-1}\,\mathrm{d}x.$$

20.15 Repeat Exercise 20.14 for the following integrals:

$$\text{(i)} \int_0^\infty \frac{\sin\sqrt{x}}{x(1+x^2)}\,\mathrm{d}x, \qquad \text{(ii)} \int_0^\infty \frac{x}{\sinh x}\,\mathrm{d}x,$$

$$\text{(iii)} \int_0^\infty \frac{\log x}{x^2+x-2}\,\mathrm{d}x, \qquad \text{(iv)} \int_0^\infty \frac{\log(1+x^2)}{x^2}\,\mathrm{d}x.$$

The remaining exercises are somewhat more challenging than the preceding ones, or involve the more advanced techniques introduced in the final section of the chapter, or both.

20.16 Prove that

$$\text{(i)} \sum_{n=1}^\infty n^{-4} = \frac{\pi^4}{90}, \qquad \text{(ii)} \sum_{n=0}^\infty (-1)^n(2n+1)^{-3} = \frac{\pi^3}{32}.$$

20.17 Prove, with the aid of Jordan's inequality, that

$$\text{(i)} \int_0^\infty \frac{x\sin ax}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2}\mathrm{e}^{-a} \quad (a > 0),$$

$$\text{(ii)} \int_0^\infty \frac{x^3\sin x}{1+x^4}\,\mathrm{d}x = \frac{\pi}{2}\mathrm{e}^{-\pi/\sqrt{2}}\cos\frac{\pi}{\sqrt{2}}.$$

20.18 Evaluate $\displaystyle\int_0^\infty \frac{\sin^3 x}{x^3}\,\mathrm{d}x$. (Hint: consider $\mathrm{e}^{3\mathrm{i}z} - 3\mathrm{e}^{\mathrm{i}z} + 2$.)

20.19 By integrating round a suitable sector, prove that

$$\int_0^\infty \mathrm{e}^{-x^2}\sin(x^2)\,\mathrm{d}x = \frac{1}{4}\sqrt{\pi}\sqrt{\sqrt{2}-1}.$$

20.20 What is wrong with the following argument? The substitution $y = (a-\mathrm{i}b)x$ gives, for $a > 0$ and $b \in \mathbb{R}$,

$$\int_0^\infty \mathrm{e}^{-ax^2}\mathrm{e}^{ibx}\,\mathrm{d}x - \frac{1}{a-\mathrm{i}b}\int_0^\infty \mathrm{e}^{-u}\,\mathrm{d}u = \frac{a+\mathrm{i}b}{a^2+b^2}.$$

Give a correct derivation, by integrating $e^{-z}$ round a suitable sector. Deduce the values of

$$\int_0^\infty e^{-ax^2} \cos bx \, dx \quad \text{and} \quad \int_0^\infty e^{-ax^2} \sin bx \, dx.$$

20.21 (This exercise contains a miscellany of integrals for which you will need to draw on all the techniques in this chapter.) Evaluate

$$\text{(i)} \int_0^\infty \frac{x^2}{(1+x^2)^2} \, dx, \qquad \text{(ii)} \int_0^\infty \frac{x - \sin x}{x^3(1+x^2)} \, dx$$

$$\text{(iii)} \int_0^\infty \frac{\sinh ax}{\sinh x} \, dx, \qquad \text{(iv)} \int_0^\infty \frac{x^{\frac{1}{2}} \log x}{(1+x)} \, dx.$$

20.22 In the plane cut along $[0, 1]$, take $(z(z-1))^{\frac{1}{2}}$ to be the holomorphic branch of the square root which is real and positive at a given point $a > 1$. Prove that

$$2i \int_0^1 \frac{1}{(x(1-x))^{\frac{1}{2}}(a-x)} \, dx = \int_\gamma \frac{1}{(z(z-1))^{\frac{1}{2}}(a-z)} \, dz,$$

where $(x(1-x))^{\frac{1}{2}}$ is positive on $[0, 1]$ and $\gamma$ is a positively oriented contour enclosing $0$ and $1$ but not enclosing $a$. Hence show that

$$\int_0^1 \frac{1}{(x(1-x))^{\frac{1}{2}}(a-x)} \, dx = \frac{\pi}{(a(a-1))^{\frac{1}{2}}}.$$

# 21 The Laplace transform

On one level, integral transforms provide a versatile and systematic method for solving equations; on another, they form the starting point for a rich theory having connections with many important branches of pure and applied mathematics. This chapter and the following one treat the rudiments of the theory and applications of the best-known and most useful integral transforms: the Laplace transform and the Fourier transform. As explained in the preface, the emphasis is on the part complex analysis plays. Readers who want to concentrate on methods, taking on trust the theory which underpins these, can simply skip the more theoretical parts of this chapter. Others will wish to understand the theory and for them we provide an introductory account of this.

The motivating idea is a very simple one: if you cannot solve a given problem, transform it into a simpler one you can solve. Find the solution of the simpler problem, and then use this to capture the solution to the original problem. In 21.20–21.22 we illustrate how the Laplace transform can be used to solve certain differential equations. In the preceding sections we set up the necessary machinery: we establish the basic properties of the transform and present methods of evaluating and inverting it.

## Basic properties and evaluation of Laplace transforms

Laplace transforms, and the Fourier transforms studied in the following chapter, are defined by integrals over unbounded intervals, $[0, \infty)$ in the first case and $\mathbb{R}$ in the second. These integrals may be taken to be Lebesgue integrals or (absolutely convergent improper) Riemann integrals. This is feasible because, in deriving theorems about transforms, the calculations are essentially the same whichever theory of integration is used. The difference comes in the way the steps are justified once suitable conditions are imposed on the functions. The following note allows us to accommodate both a Lebesgue and a Riemann approach to integration in this chapter and the next.

**21.1 Technical note.** Let $J$ denote either $[0, \infty)$ or $\mathbb{R}$.

For a Lebesgue integral treatment, $\mathcal{I}(J)$ should be intepreted as $\mathrm{L}^1(J)$, the (complex-valued) Lebesgue integrable functions on $J$.

For a Riemann integral treatment, $\mathcal{I}(J)$ should be interpreted as the set of (complex-valued) functions on $J$ which are piecewise continuous on any closed bounded subinterval of $J$ and such that each of $f$ and $|f|$ has an improper integral on $J$. The terms used here were defined in 10.2 and 19.14. These conditions could be relaxed somewhat; see **[1]**, Chapter 15.

**21.2 Introducing the Laplace transform.** The Laplace transform of a complex-valued function $f$ defined on $[0, \infty)$ is given by

$$\overline{f}(p) := \int_0^\infty f(t)\mathrm{e}^{-pt}\,\mathrm{d}t,$$

if $f(t)\mathrm{e}^{-pt} \in \mathcal{I}([0,\infty))$. Here $p$ is allowed to be complex. We follow convention in using $p$ rather than $z$ to denote the variable.

We may think of the Laplace transform as an operator $\mathfrak{L}$ taking a function $f$ to its transform $\overline{f}$, so that $\mathfrak{L}f = \overline{f}$. We shall allow an abuse of notation and insert or omit the variables ($t$ and $p$) as expedient. Specifically, it will be convenient to adopt as alternative notations for the transform of $f$ both $\overline{f}(p)$ and $\mathfrak{L}[f(t)]$. We usually use the latter style for the transform of a concrete function and the former when writing the resulting function of $p$. Thus, for $f(t) = t\sin t$ we find that $\overline{f}(p) = 2p(p^2+1)^{-2}$, and also denote this function by $\mathfrak{L}[t\sin t]$ (see 21.8).

To give an impression of how functions transform under $f \mapsto \overline{f}$ we immediately present some examples.

**21.3 Elementary examples.** Direct integration, combined with 21.5–21.7 below, enables a catalogue of basic transforms to be constructed. We record some of the most useful.

**Table 21.1** Some Laplace transforms

| $f(t)$ | $\overline{f}(p)$ | valid for |
|---|---|---|
| $1$ | $1/p$ | $\operatorname{Re} p > 0$ |
| $t^n \ (n = 1, 2, \dots)$ | $n!/p^{n+1}$ | $\operatorname{Re} p > 0$ |
| $\mathrm{e}^{-at}$ | $1/(p+a)$ | $\operatorname{Re} p > -\operatorname{Re} a$ |
| $\cos\omega t$ | $p/(p^2+\omega^2)$ | $\operatorname{Re} p > \lvert\operatorname{Im}\omega\rvert$ |
| $\sin\omega t$ | $\omega/(p^2+\omega^2)$ | $\operatorname{Re} p > \lvert\operatorname{Im}\omega\rvert$ |

**21.4 The existence of the Laplace transform.** In order that the integral defining $\overline{f}$ should exist, it is certainly necessary that $f$ should be a reasonably well-behaved function. The table above signals that we may also expect to need a restriction on $p$.

In practice, we are principally interested in differentiable functions, and for such functions, or more generally for continuous functions, the integral of $f(t)\mathrm{e}^{-pt}$ exists over any bounded interval $[0, R]$, for any $p$. The exponential factor is an asset in ensuring good behaviour as $R \to \infty$. Note that $|\mathrm{e}^{-pt}| = \mathrm{e}^{-\operatorname{Re} pt}$ and that, for $\operatorname{Re} p > 0$, this exponential decays rapidly as $t \to \infty$. In particular, if $f$ is continuous on $[0, \infty)$ and such that $|f(t)| \leqslant M\mathrm{e}^{-ct}$ for $t \geqslant T$, where $M$ and $T$ are constants, then $\overline{f}(p)$ exists for $\operatorname{Re} > c$. On the other hand, a function such as $\mathrm{e}^{\mathrm{e}^t}$ which grows exceptionally rapidly will not have a Laplace transform.

We now present some general results which are frequently useful. Elementary calculations yield the following lemma.

**21.5 New transforms from old.** Provided the transforms involved exist:

(1) $\mathfrak{L}[af(t) + bg(t)] = a\mathfrak{L}[f(t)] + b\mathfrak{L}[g(t)]$ for any constants $a$ and $b$ in $\mathbb{C}$ (that is, $\mathfrak{L}$ is linear);

(2) $\mathfrak{L}[f(t/a)] = a\overline{f}(pa)$ $(a > 0)$;

(3) $\mathfrak{L}[\mathrm{e}^{-at}f(t)] = \overline{f}(p + a)$ $(a \in \mathbb{C})$;

(4) $\mathfrak{L}[f(t - a)H(t - a)] = \mathrm{e}^{-ap}\overline{f}(p)$, where $H$ is the Heaviside function defined by $H(t) = 1$ $(t \geqslant 0)$, $H(t) = 0$ $(t < 0)$ (that is, $H = \chi_{[0,\infty)}$).

**21.6 The Laplace transform of a derivative.** Under appropriate conditions on $f$,

$$\mathfrak{L}[f^{(n)}(t)] = p^n\overline{f}(p) - p^{n-1}f(0) - \cdots - f^{(n-1)}(0) \qquad (n = 1, 2, \ldots).$$

To derive the formula, we integrate by parts to obtain

$$\begin{aligned}\mathfrak{L}[f^{(n)}(t)] &= \Big[f^{(n-1)}(t)\mathrm{e}^{-pt}\Big]_0^\infty + p\int_0^\infty f^{(n-1)}(t)\mathrm{e}^{-pt}\,\mathrm{d}t \\ &= -f^{(n-1)}(0) + p\mathfrak{L}[f^{(n-1)}(t)],\end{aligned}$$

provided $f^{(n-1)}(t)\mathrm{e}^{-pt} \to 0$ as $t \to \infty$. We then repeat the process or, more formally, use induction. Sufficient conditions for the formula to be valid are

(i) $f', \ldots, f^{(n)}$ all exist and the transforms of $f, f', \ldots, f^{(n)}$ all exist,

(ii) $f^{(n)}$ is continuous on $[0, \infty)$, and
(iii) $f^{(k)}(t)e^{-pt} \to 0$ as $t \to \infty$ $(k = 0, \ldots, n-1)$.

Notice how the operator $\mathfrak{L}$ converts derivatives of $f$ into algebraic expressions involving $\overline{f}$. This is the crucial property involved in the solution of differential equations by transform methods.

**21.7 The derivative of a Laplace transform.** Suppose $\overline{f}(p)$ exists for $\operatorname{Re} p > c$. Then $\overline{f}(p)$ is holomorphic for $\operatorname{Re} p > c$, with derivatives given by differentiation under the integral sign, so that

$$\mathfrak{L}[t^n f(t)] = (-1)^n \left(\frac{\mathrm{d}}{\mathrm{d}p}\right)^n \overline{f}(p).$$

**Outline proof** Fix $p$ such that $\operatorname{Re} p > c$ and write $\operatorname{Re} p - c = 2\eta$. Let $h$ be such that $|h| < \eta$, so that $\operatorname{Re}(p+h) > c + \eta$. Then

$$\begin{aligned}\left|\frac{\overline{f}(p+h) - \overline{f}(p)}{h} + \int_0^\infty t f(t)e^{-pt}\,\mathrm{d}t\right| &= \left|\int_0^\infty f(t)e^{-pt}\left(\frac{e^{-ht}-1}{h} + t\right)\mathrm{d}t\right| \\ &= \left|\int_0^\infty f(t)e^{-pt}\sum_{n=2}^\infty \frac{(th)^n}{n!h}\,\mathrm{d}t\right| \\ &\qquad \text{(using the expansion for } e^{-ht}\text{)} \\ &\leqslant |h| \int_0^\infty \left|f(t)e^{-pt}\right| t^2 e^{t|h|}\,\mathrm{d}t \\ &\leqslant |h| \int_0^\infty \left|f(t)e^{-ct}t^2e^{-\eta t}\right|\mathrm{d}t.\end{aligned}$$

This tends to zero as $h \to 0$, since $t^2e^{-\eta t}$ is bounded on $[0, \infty)$ and $|f(t)e^{-ct}|$ is integrable by hypothesis. This gives the existence of $\mathfrak{L}[f'(t)]$ and the required formula for it. Higher-order derivatives are handled in the same way. □

## Inversion of Laplace transforms

We next consider how a function $f$ can be recovered from its Laplace transform $\overline{f}$. Thus we need to ask whether, given a function $g(p)$, we can find $f$ such that $\overline{f} = g$. The simplest method is obviously 'inspection': recognizing a function with the required transform, with the aid of 21.5–21.3.

## 21.8 Inversion of Laplace transforms: elementary examples.

- $\dfrac{p+1}{p^2(p-1)} = \dfrac{2}{p-1} - \dfrac{1+2p}{p^2} = \mathfrak{L}(2e^t - t - 2)$ (by 21.5(1) & (3) and 21.3).
- $\dfrac{1}{p^2+4p+20} = \dfrac{1}{(p+2)^2+16} = \mathfrak{L}\left(\frac{1}{4}e^{-2t}\sin 4t\right)$ (by 21.5(1) & (3) and 21.3).
- Consider

$$\begin{aligned}\frac{p}{(p^2+\omega^2)^2} &= \frac{1}{4i\omega}\left(\frac{1}{(p-i\omega)^2} - \frac{1}{(p+i\omega)^2}\right)\\ &= \frac{1}{4i\omega}\left(\mathfrak{L}(te^{i\omega t}) - \mathfrak{L}(te^{-i\omega t})\right) && \text{(by 21.5 and 21.3 or 21.7)}\\ &= \mathfrak{L}\left(\frac{t\sin\omega t}{2\omega}\right) && \text{(by 21.5(1)).}\end{aligned}$$

Alternatively,

$$\frac{p}{(p^2+\omega^2)^2} = \frac{1}{2}\frac{d}{dp}\left(-\frac{1}{p^2+\omega^2}\right) = \mathfrak{L}\left(\frac{t\sin\omega t}{2\omega}\right) \qquad \text{(by 21.7 and 21.3).}$$

See 21.12 and 21.15 for further methods for obtaining the same result.

Without recourse to a published table of transforms or to a computer algebra package, the inspection method is of limited use. Fortunately there is an inversion theorem which applies to a very wide range of functions (and which is used in the compilation of extensive tables). The version of the theorem we give is not the most general. The smoothness condition we impose is no hindrance in applications and facilitates the proof. However this is still highly technical. We indicate in 22.9 how the theorem follows from a corresponding theorem for Fourier transforms (stated in 22.8, proved in [**6**]).

**21.9 Piecewise smooth functions.** We introduced piecewise continuous functions in Chapter 11. We note now that if $f$ is piecewise continuous and is defined on an open interval containing $t$, then the left-hand and right-hand limits, $f(t-)$ and $f(t+)$, exist. Jump discontinuities occur at those points where $f(t-) \neq f(t+)$.

Let $J = \mathbb{R}$ or $[0, \infty)$. We say that a (real- or complex-valued) function $f$ is **piecewise smooth on** $J$ if $f$ and $f'$ are piecewise continuous on every closed bounded subinterval of $J$. This definition may seem daunting. In fact, piecewise smooth functions arise frequently and are easily recognized; see Fig. 21.1 for

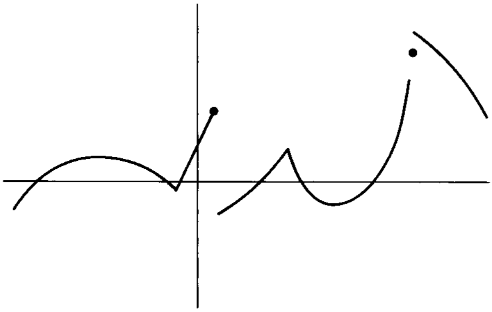


**Figure 21.1** A piecewise smooth function

an archetypal example. Any continuously differentiable function is, of course, piecewise smooth.

**21.10 Inversion theorem for the Laplace transform.** Suppose that $f$ is piecewise smooth on $[0, \infty)$ and that $\overline{f}(p)$ exists for $\operatorname{Re} p > c \geqslant 0$. Then, for $t > 0$,

$$\frac{1}{2}\big(f(t+) + f(t-)\big) = \frac{1}{2\pi \mathrm{i}} \lim_{R \to \infty} \int_{\sigma - \mathrm{i}R}^{\sigma + \mathrm{i}R} \overline{f}(p) \mathrm{e}^{pt} \, \mathrm{d}p \quad (\sigma > c).$$

The left-hand side simplifies to $f(t)$ if $f$ is continuous at $t$. The integral on the right-hand side is along the vertical line segment $[\sigma - \mathrm{i}R, \sigma + \mathrm{i}R]$ in $\mathbb{C}$; it is independent of the value of $\sigma$ $(> c)$.

The inversion theorem guarantees that any continuous and piecewise smooth function is uniquely determined by its transform. It can in fact be proved that, for a function $f$ which is merely continuous, $\overline{f} \equiv 0$ implies $f \equiv 0$ (Lerch's theorem). This uniqueness property is tacitly used whenever inverse transforms are obtained by inspection.

The inversion integral can frequently be evaluated by contour integration, using the techniques developed in Chapters 19 and 20. The following lemma gives a handy sufficient condition for an inverse transform to be a sum of residues.

**21.11 Lemma (inverse Laplace transform via Cauchy's residue theorem).** Let $g$ be holomorphic except for a finite number of poles at $a_1, \ldots, a_n$, and suppose that there exist constants $M$ and $k$ such that

$$|g(p)| \leqslant M \, |p|^{-k} \quad \text{for large } |p| \, .$$

Then, for $t > 0$ and $\sigma > \operatorname{Re} a_j$ $(j = 1, \ldots, n)$,

$$\frac{1}{2\pi i} \lim_{R\to\infty} \int_{\sigma - iR}^{\sigma + iR} g(p)e^{pt}\, dp = \sum_{j=1}^{n} \operatorname{res}\left\{g(p)e^{pt}; a_j\right\}.$$

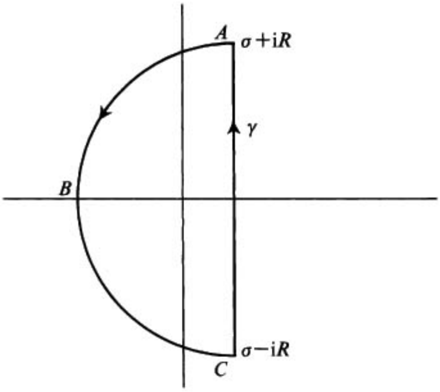



**Figure 21.2** Contour for Lemma 21.11

**Proof** We integrate $g(p)e^{pt}$ round the semicircular contour $\gamma$ shown in Fig. 21.2 and apply Cauchy's residue theorem. On the semicircular arc ABC, $p = \sigma + iRe^{i\theta}$ $(\theta \in [\pi/2, 3\pi/2])$ and $|p| \geqslant R - \sigma$ (by 1.9(3)). The given bound on $g$ gives, for large $R$,

$$\begin{aligned} \left| \int_{ABC} g(p)e^{pt}\, dp \right| &\leqslant \int_{\pi/2}^{3\pi/2} M\,|R-\sigma|^{-k}\, |e^{(\sigma + iRe^{i\theta})t} Rie^{i\theta}|\, d\theta \\ &\leqslant \int_{\pi/2}^{3\pi/2} M\,|R-\sigma|^{-k}\, e^{\sigma t - tR\cos\theta} R\, d\theta \\ &= 2\int_{0}^{\pi/2} M\,|R-\sigma|^{-k}\, e^{\sigma t - tR\sin\varphi}\, d\varphi \end{aligned}$$

(putting $\varphi = \theta - \pi/2$). The final integral tends to 0 as $R \to \infty$ (Jordan's inequality is needed if $0 < k \leqslant 1$; see 19.7). □

**21.12 Examples (Laplace transform inversion theorem).**

- Consider again the function

$$g(p) = \frac{p+1}{p^2(p-1)};$$

this is holomorphic except for a simple pole at 1 and a double pole at 0, and it satisfies the conditions for Lemma 21.11. Hence $g = \overline{f}$, where

$$f(t) = \operatorname{res}\left\{\frac{(p+1)e^{pt}}{p^2(p-1)};0\right\} + \operatorname{res}\left\{\frac{(p+1)e^{pt}}{p^2(p-1)};1\right\} = 2e^t - 2 - t.$$

Note that use of the Laplace inversion theorem avoids the partial fraction decomposition employed in 21.8.

- By Lemma 21.11,

$$\begin{aligned} p(p^2+\omega^2)^{-2} &= \operatorname{res}\left\{\frac{pe^{pt}}{(p^2+\omega^2)^2}; i\omega\right\} + \operatorname{res}\left\{\frac{pe^{pt}}{(p^2+\omega^2)^2}; -i\omega\right\} \\ &= \left[\frac{d}{dp}\left(\frac{pe^{pt}}{(p+i\omega)^2}\right)\right]_{p=i\omega} + \left[\frac{d}{dp}\left(\frac{pe^{pt}}{(p-i\omega)^2}\right)\right]_{p=-i\omega}, \end{aligned}$$

and this, by a straightforward calculation, is $(t\sin\omega t)/(2\omega)$, as we expect from 21.8.

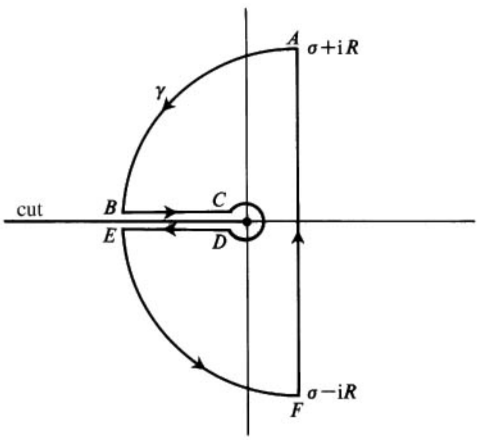



**Figure 21.3** Contour used in calculating $\mathfrak{L}^{-1}(\sqrt{p})$

- We find a function with Laplace transform (a holomorphic branch of) $1/\sqrt{p}$. We cut the plane along the negative real axis and take $p = |p|\,e^{i\theta}$ $(-\pi <$

$\theta \leqslant \pi$) and $\sqrt{p} = |p|^{1/2} e^{i\theta/2}$. We wish to evaluate

$$\frac{1}{2\pi i} \lim_{R\to\infty} \int_{\sigma - iR}^{\sigma + iR} \frac{1}{\sqrt{p}} e^{pt} \, dp \quad (\sigma > 0).$$

We use the keyhole contour $\gamma$ shown in Fig. 21.3. By Cauchy's theorem, the integral of $e^{pt}/\sqrt{p}$ round $\gamma$ is zero. We can prove that the integrals along $AB$ and $EF$ tend to zero, in just the same way as in 21.11. Also

$$\left| \int_{\gamma(0;\varepsilon)} \frac{1}{\sqrt{p}} e^{pt} \, dp \right| \leqslant \int_0^{2\pi} \frac{1}{\sqrt{\varepsilon}} e^{\varepsilon t \cos\theta} \varepsilon \, d\theta = \mathbf{O}(\varepsilon^{\frac{1}{2}}).$$

On $BC$, $\sqrt{p} = i\sqrt{x}$ $(x > 0)$ and on $DE$, $\sqrt{p} = -i\sqrt{x}$ $(x > 0)$. The integrals along $BC$ and $DE$ combine to give

$$-2\int_\varepsilon^R \frac{1}{i\sqrt{x}} e^{-xt} \, dx.$$

Hence, letting $\varepsilon \to 0$ and $R \to \infty$,

$$\begin{aligned} \mathfrak{L}^{-1}(1/\sqrt{p}) &= \frac{1}{\pi} \int_0^\infty \frac{1}{\sqrt{x}} e^{-xt} \, dx \\ &= \frac{2}{\pi\sqrt{t}} \int_0^\infty e^{-y^2} \, dy \qquad \text{(putting } xt = y^2\text{)} \\ &= 1/\sqrt{\pi t} \qquad \text{(by 20.15).} \end{aligned}$$

The following two results extend the range of functions for which we can calculate the inverse Laplace transform. Under suitable conditions, thr first allows us to invert a product of two transforms, and the second to invert term-by-term a transform $\overline{f}(p)$ expressed as a series of negative powers of $p$.

**21.13 Convolution theorem for the Laplace transform.** Suppose $f$ and $g$ are such that $\overline{f}$ and $\overline{g}$ exist for $\operatorname{Re} p > c$. Then $\overline{f}\,\overline{g} = \overline{h}$ for $\operatorname{Re} p > c$, where $h$ is the **convolution** of $f$ and $g$ defined by

$$h(y) = \int_0^y f(t) g(y - t) \, dt \quad (y \geqslant 0).$$

**Informal derivation** Consider

$$\begin{aligned}
\overline{f}(p)\overline{g}(p) &= \int_0^\infty f(t)\mathrm{e}^{-pt}\,\mathrm{d}t \int_0^\infty g(s)\mathrm{e}^{-ps}\,\mathrm{d}s \\
&= \int_0^\infty \int_0^\infty f(t)g(s)\mathrm{e}^{-p(t+s)}\,\mathrm{d}s\,\mathrm{d}t \\
&= \int_0^\infty \int_t^\infty f(t)g(y-t)\mathrm{e}^{-py}\,\mathrm{d}y\,\mathrm{d}t \qquad \text{(putting } y = s+t \text{ (with } t \text{ fixed))} \\
&= \int_0^\infty \left\{ \int_0^y f(t)g(y-t)\,\mathrm{d}t \right\} \mathrm{e}^{-py}\,\mathrm{d}y \\
&= \overline{h}(p).
\end{aligned}$$

The penultimate line is obtained by switching the order in which the two integrations are performed, noting that $t \in [0, y]$ if and only if $y \in [t, \infty)$ (for $t, y \geqslant 0$). It is this interchange that needs justification.

**21.14 Theorem (inverting a series expansion).** Suppose that the continuous function $f$ satisfies the conditions for the inversion theorem (21.10) and that $\overline{f}(p)$ is expressible as

$$\overline{f}(p) = \sum_{n=0}^{\infty} a_n p^{-n-1},$$

where the series on the right-hand side converges for $|p| > \rho$. Then, for $t > 0$,

$$f(t) = \sum_{n=0}^{\infty} \frac{a_n}{n!} t^n$$

(that is, term-by-term inversion is legitimate).

**Outline proof** Since $g(p) = \sum_{n=0}^{\infty} a_n p^{-n-1}$ converges for $|p| > \rho$, 6.1(3) implies that $\sum |a_n| \, r^{-n-1}$ converges for $r > \rho$. Thus

(i) $g(p)$ is holomorphic for $|p| > \rho$ (cf. 6.8);

(ii) for $|p| \geqslant S > \rho$, $|g(p)| \leqslant |p|^{-1} \sum_{n=0}^{\infty} |a_n| \, S^{-n} = \mathbf{O}(|p|^{-1})$;

(iii) on $|p| = S > \rho$, $|\mathrm{e}^{pt}|$ is bounded, by $K$ say, so that

$$\left| a_n p^{-n-1} \mathrm{e}^{pt} \right| \leqslant M_n := K \, |a_n| \, S^{-n-1}$$

and $\sum M_n$ converges.

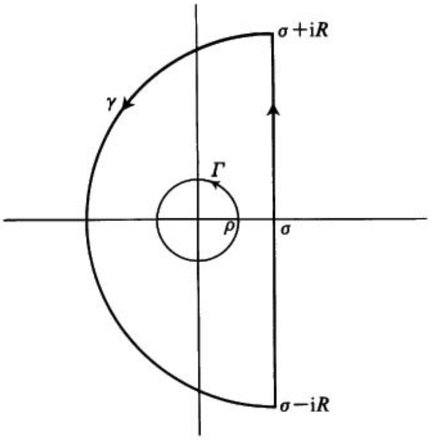



**Figure 21.4** Contours for the proof of Theorem 21.14

By the inversion theorem we have, for $\sigma > \rho$,

$$\begin{aligned} f(t) &= \frac{1}{2\pi \mathrm{i}} \lim_{R\to\infty} \int_{\sigma-\mathrm{i}R}^{\sigma+\mathrm{i}R} g(p)\mathrm{e}^{pt}\,\mathrm{d}p \\ &= \frac{1}{2\pi \mathrm{i}} \lim_{R\to\infty} \int_{\gamma} g(p)\mathrm{e}^{pt}\,\mathrm{d}p && \text{(taking } \gamma \text{ as in Fig. 21.4 and using 21.11, (i), (ii))} \\ &= \frac{1}{2\pi \mathrm{i}} \int_{\gamma(0;S)} g(p)\mathrm{e}^{pt}\,\mathrm{d}p && \text{(using (ii) and 11.9; here } \rho < S < \sigma) \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi\mathrm{i}} a_n \int_{\gamma(0;S)} p^{-n-1}\mathrm{e}^{pt}\,\mathrm{d}p && \text{(using (iii) and 14.2)} \\ &= \sum_{n=0}^{\infty} a_n t^n/n! && \text{(see Table 21.1).} \qquad \square \end{aligned}$$

**21.15 Examples (inversion via the Convolution theorem and via series expansion).** Consider once again $p(p^2+\omega^2)^{-2}$.

- By 21.13 and elementary trigonometry,

$$\begin{aligned} \frac{p}{(p^2+\omega^2)^2} &= \frac{1}{\omega}\mathfrak{L}(\cos\omega t)\mathfrak{L}(\sin\omega t) \\ &= \frac{1}{\omega}\mathfrak{L}\left(\int_0^t \cos\omega y \sin(\omega(t-y))\,\mathrm{d}y\right) \end{aligned}$$

$$= \frac{1}{2\omega}\mathfrak{L}\left(\int_0^t \sin\omega(y + (t-y)) - \sin\omega(y - (t-y))\,\mathrm{d}y\right)$$
$$= \mathfrak{L}\left(\frac{t\sin\omega t}{2\omega}\right).$$

- We have

$$\begin{aligned}\frac{p}{(p^2+\omega^2)^2} &= \sum_{n=0}^{\infty}(-1)^n(n+1)\frac{\omega^{2n}}{p^{2n+3}} \qquad (|p| > |\omega|)\\ &= \frac{1}{2\omega}\mathfrak{L}\left(t\sum_{n=0}^{\infty}(-1)^n\frac{(\omega t)^{2n+1}}{(2n+1)!}\right) \qquad \text{(by 21.14)}\\ &= \mathfrak{L}\left(\frac{t\sin\omega t}{2\omega}\right).\end{aligned}$$

**21.16 Inversion of the Laplace transform: summary.** We now have a number of inversion methods at our disposal:

- inspection, using as a starting point the transforms in Table 21.1 and exploiting the results in 21.5;
- direct computation of the inversion integral, usually with the aid of Lemma 21.11;
- use of the Convolution theorem, 21.13, to invert a function recognizable as a product of known transforms or of more easily inverted functions;
- series expansion followed by term-by-term inversion based on Theorem 21.14.

As we have seen in 21.8 (21.12) and 21.15, there is often a choice of viable methods.

## Applications

In this section we illustrate the use of the Laplace transform in the solution of simple ordinary and partial differential equations and integral equations. We start with a simple integral equation.

**21.17 Example (a Volterra integral equation).** We show how the Convolution theorem can be used to solve the integral equation

$$m(t) = (1 - e^{-\lambda t}) + \lambda \int_0^t m(t-x)e^{-\lambda x}\, dx,$$

where $\lambda$ is a positive constant. We apply $\mathfrak{L}$ and use the Convolution theorem. This gives

$$\overline{m}(p) = \frac{1}{p} - \frac{1}{\lambda + p} + \lambda \frac{\overline{m}(p)}{\lambda + p}.$$

Hence $\overline{m}(p) = \lambda/p^2$, so that $m(t) = \lambda t$ provides the required solution.

**21.18 Solving differential equations using the Laplace transform.** We reveal the idea behind the method with a very simple example, taken from electrical circuit theory. Consider the problem of finding the function $I(t)$, for $t \geqslant 0$, which satisfies the equation

$$L\frac{dI}{dt} + RI = E \quad \text{with } I = 0 \text{ when } t = 0;$$

where $L$, $R$, $E$ are constants. We multiply the given equation by $e^{-pt}$ and integrate over $[0, \infty)$ to get (by linearity of $\mathfrak{L}$),

$$L\int_0^\infty \frac{dI}{dt} e^{-pt}\, dt + R\int_0^\infty I(t)e^{-pt}\, dt = E\int_0^\infty e^{-pt}\, dt.$$

By 21.6 the first integral is $p\overline{I}(p) - I(0)$. Hence, using the given initial condition, we have

$$Lp\overline{I}(p) + R\overline{I}(p) = \frac{E}{p}.$$

Therefore

$$\overline{I}(p) = \frac{E}{p(R + Lp)} = \frac{E}{R}\left(\frac{1}{p} - \frac{1}{p + R/L}\right).$$

We recognize the right-hand side as the transform of $\dfrac{E}{R}(1 - e^{-Rt/L})$. This function is the unique solution for $I(t)$.

We can apply the same idea much more widely. Suppose $t$ takes values in $[0, \infty)$. Then the Laplace transform will convert

- an ordinary differential equation for $f(t)$ with constant coefficients into an algebraic equation for $\overline{f}(p)$ (see the example above and Example 21.20);
- an ordinary differential equation for $f(t)$ with polynomial coefficients into an ordinary differential equation for $\overline{f}(p)$ (Example 21.21);

- simultaneous ordinary differential equations, with constant coefficients, in $f_1(t), \ldots, f_n(t)$ into simultaneous equations in $\overline{f_1}(p), \ldots, \overline{f_n}(p)$ (Exercise 21.11);
- a partial differential equation for $u(x,t)$ (of suitable type) into an ordinary differential equation for $\overline{u}(x,p)$ with variable $x$ (Examples 21.22 and Exercise 21.17).

The statements above should not be treated as rules which are universally applicable, but rather as guidance on what form transformed problems will take. If the method is to be carried through successfully to solve a given problem, certain technical provisos must be added and we must include boundary and/or initial conditions, as needed. For the Laplace transform, we need one independent variable, $t$, which has domain $[0, \infty)$; for physical systems evolving in time, we are likely to take $t$ to be time. Where we have more than one variable defined on $[0, \infty)$ the form of the initial or boundary conditions often determines on which variable we should operate by $\mathfrak{L}$.

Transform methods are particularly suitable where the equations involve unknown constants or functions, often standing for physical quantities. Equations involving only numerical constants and given functions are usually more easily solved by more elementary means.

**21.19 Important tactical tips.** Before we embark on examples we amplify a comment made earlier. It is often relatively easy to *obtain* a transform solution, but if we do not establish the validity of each step in the calculation we cannot be sure that our 'solution' is correct. The following remarks should help to prevent an excessively pedantic or an excessively cavalier attitude.

- ⊙ **Don't break the rules** For example, don't write down integrals of functions which are not integrable. Obey the rules when using contour integration: for example, integrals round arcs must tend to the limits claimed and a cut plane must be used with a multifunction.
- ⊙ **Amenable functions** Serious problems are likely to arise only with 'pathological' functions: solutions of differential equations are, a priori, reasonably smooth, and hence behave well under transforms. However it may be necessary to impose restrictions on a function's rate of growth and limiting behaviour which are not built into the original problem. In many applied problems such technical restrictions are acceptable on physical grounds.
- ⊙ **Don't make work!** In many cases delicate analysis is neither necessary nor appropriate. It is frequently easier to carry through the calculations without checking the validity of the individual steps, and to verify at the

end that the 'solution' found does indeed satisfy the given equation and any additional conditions. Some familiarity with existence and uniqueness theorems is valuable here for knowing what to expect by way of solutions.

**21.20 Example (a linear ordinary differential equation).** Consider

$$f''(t) + f(t) = \begin{cases} \cos t & (0 \leqslant t \leqslant \pi), \\ 0 & (t > \pi), \end{cases}$$

given $f(0) = f'(0) = 0$. We operate on the differential equation by $\mathfrak{L}$ to get

$$\mathfrak{L}[f''(t)] + \mathfrak{L}[f(t)] = \int_0^\pi \cos t \, e^{-pt} \, dt = \frac{p(1 - e^{-p\pi})}{1 + p^2}.$$

Hence, using 21.6 and the initial conditions,

$$(1 + p^2)\bar{f}(p) = \frac{p(1 + e^{-p\pi})}{1 + p^2}.$$

We have shown in many different ways that $p(p^2 + 1)^{-2}$ is the Laplace transform of $\frac{1}{2}t \sin t$. Hence, by 21.5(4),

$$\begin{aligned} f(t) &= \frac{1}{2}t \sin t + \frac{1}{2}(t - \pi)\sin(t - \pi)H(t - \pi) \\ &= \begin{cases} \frac{1}{2}t \sin t & (0 \leqslant t \leqslant \pi), \\ \frac{1}{2}\pi \sin t & (t > \pi). \end{cases} \end{aligned}$$

We may consider, more generally,

$$f''(t) + f(t) = k(t) \quad (t \geqslant 0), \qquad \text{subject to } f(0) = f'(0) = 0.$$

The transformed equation is now

$$\bar{f}(p) = \frac{\bar{k}(p)}{1 + p^2}.$$

Hence, for any reasonably well-behaved function $k$,

$$f(t) = \int_0^t k(x) \sin(t - x) \, dx \quad (t \geqslant 0),$$

by the Convolution theorem.

**21.21 Example (an ordinary differential equation with polynomial coefficients).** We find a solution of Bessel's equation of order zero:

$$tf''(t) + f'(t) + tf(t) = 0.$$

We first apply $\mathfrak{L}$ and use 21.6 and 21.7. The equation transforms into

$$\frac{\mathrm{d}}{\mathrm{d}p}\left(p^2\overline{f}(p) - pf(0) - f(0)\right) + \left(p\overline{f}(p) - f(0)\right) - \frac{\mathrm{d}}{\mathrm{d}p}\overline{f}(p) = 0.$$

Hence

$$(p^2 + 1)\frac{\mathrm{d}}{\mathrm{d}p}\overline{f}(p) = \overline{f}(p).$$

This is satisfied by

$$\overline{f}(p) = A(p^2 + 1)^{-\frac{1}{2}} \quad (A \text{ constant}),$$

where the right-hand side is defined to be holomorphic in the plane cut between $\mathrm{i}$ and $-\mathrm{i}$ (see 9.2). For $|p| > 1$,

$$(p^2 + 1)^{-\frac{1}{2}} = \sum_{n=0}^{\infty}(-1)^n \frac{(2n)!}{(n!)^2 2^{2n} p^{2n+1}}.$$

Theorem 21.14 now implies that

$$f(t) = A\sum_{n=0}^{\infty}(-1)^n \frac{1}{(n!)^2}\left(\frac{t}{2}\right)^{2n} = A\mathbf{J}_0(t).$$

It can be checked, with the aid of Theorem 6.8, that this series does satisfiy Bessel's equation.

Bessel's equation is of second order. It does have, as would be expected, two linearly independent solutions, $\mathbf{J}_0(t)$ and $\mathbf{Y}_0(t)$. However only $\mathbf{J}_0(t)$ is sufficiently well behaved to be obtained via the Laplace transform; $\mathbf{Y}_0(t)$ 'blows up' at $t = 0$.

### Tactical tip

- ☉ The Laplace transform method cannot find a solution of an equation if that solution does not possess a Laplace transform. Where, for a given equation, the number of independent solutions found is fewer than an existence theorem predicts, any missing solution is too ill-behaved for this method to detect it.

**21.22 Example (the diffusion equation).** We find the function $u(x,t)$ which is defined and continuous in $\{(x,t) : x \geqslant 0, t \geqslant 0\}$ and which satisfies

(i) $\dfrac{\partial u}{\partial t} = \kappa \dfrac{\partial^2 u}{\partial x^2}$ ($\kappa$ a constant $> 0$);

(ii) $u(x,0) = 0$ $(x \geqslant 0)$;

(iii) $u(0,t) = U$ $(t \geqslant 0)$, where $U$ is a constant;

(iv) $u(x,t)$ remains bounded as $x \to \infty$.

We operate by $\mathfrak{L}$ on the variable $t$, writing

$$\overline{u}(x,p) = \int_0^\infty u(x,t)\mathrm{e}^{-pt}\,\mathrm{d}t.$$

By 21.6 and (ii),

$$\int_0^\infty \frac{\partial}{\partial t} u(x,t)\mathrm{e}^{-pt}\,\mathrm{d}t = p\overline{u}(x,p).$$

Treating $p$ as fixed and assuming differentiation under the integral sign with respect to $x$ is permissible twice, (i) transforms into

$$p\overline{u}(x,p) = \kappa \frac{\mathrm{d}^2\overline{u}}{\mathrm{d}x^2}.$$

This has solution

$$\overline{u}(x,p) = A(p)\mathrm{e}^{x\sqrt{p/\kappa}} + B(p)\mathrm{e}^{-x\sqrt{p/\kappa}},$$

where $A(p)$ and $B(p)$ are functions of $p$. Here we assume the plane cut along $(-\infty, 0]$ and take, for $p = r\mathrm{e}^{\mathrm{i}\theta}$, the branch $\sqrt{p} = r^{1/2}\mathrm{e}^{\mathrm{i}\theta/2}$, where $-\pi < \theta \leqslant \pi$. For this choice, $\operatorname{Re}\sqrt{p} \geqslant 0$.

Now operate by $\mathfrak{L}$ on (iii) and (iv). From (iii),

$$\overline{u}(0,p) = U/p \quad (\operatorname{Re} p > 0),$$

while (iv) implies that $\overline{u}(x,p)$ remains bounded as $x \to \infty$. Hence $A(p) = 0$ and $B(p) = U/p$, so

$$\overline{u}(x,p) = \frac{U}{p}\mathrm{e}^{-x\sqrt{p/\kappa}} \quad (\operatorname{Re} p > 0).$$

The inversion theorem now gives, for $t > 0$,

$$u(x,t) = \frac{1}{2\pi\mathrm{i}} \lim_{R\to\infty} \int_{\sigma-\mathrm{i}R}^{\sigma+\mathrm{i}R} \frac{U}{p}\mathrm{e}^{pt-x\sqrt{p/\kappa}}\,\mathrm{d}p \quad (\sigma > 0).$$

To evaluate this, we use the keyhole contour shown in Fig. 21.3. On $BC$, $\sqrt{p} = \mathrm{i}\sqrt{r}$ and on $DE$, $\sqrt{p} = -\mathrm{i}\sqrt{r}$ $(r > 0)$. For $p = r\mathrm{e}^{\mathrm{i}\theta}$ $(-\pi < \theta \leqslant \pi)$,

$$\left|\frac{1}{p}\mathrm{e}^{-x\sqrt{p/\kappa}}\right| = \frac{1}{r}\mathrm{e}^{-x\sqrt{r/\kappa}\cos(\theta/2)} = \mathbf{O}(r^{-1}).$$

Hence, by Jordan's inequality (as in the proof of Lemma 21.11), the integrals along $AB$ and $EF$ tend to zero as $R \to \infty$. On the small circle of radius $\varepsilon$,

$$\frac{1}{p}\mathrm{e}^{pt-x\sqrt{p/\kappa}} = \frac{1}{p} + h(p),$$

where $h(p) = \mathbf{O}(1/\sqrt{\varepsilon})$. Hence

$$\int_{\gamma(0;\varepsilon)} \frac{1}{p}\mathrm{e}^{pt-x\sqrt{p/\kappa}}\,\mathrm{d}p = 2\pi\mathrm{i} + \mathbf{O}(\sqrt{\varepsilon}) \qquad \text{(by 10.4 and 10.10).}$$

The function $(U/p)\mathrm{e}^{pt-x\sqrt{p/\kappa}}$ is holomorphic inside and on the keyhole. Applying Cauchy's theorem and letting $\varepsilon \to 0$ and $R \to \infty$, we get

$$\begin{aligned} u(x,t) &= U - \frac{U}{2\pi\mathrm{i}}\int_0^\infty \mathrm{e}^{-rt+\mathrm{i}x\sqrt{r/\kappa}}\frac{1}{r}\,\mathrm{d}r + \frac{U}{2\pi\mathrm{i}}\int_0^\infty \mathrm{e}^{-rt-\mathrm{i}x\sqrt{r/\kappa}}\frac{1}{r}\,\mathrm{d}r \\ &= U - \frac{U}{\pi}\int_0^\infty \sin\left(x\sqrt{r/\kappa}\right)\mathrm{e}^{-rt}\frac{1}{r}\,\mathrm{d}r \\ &= U - \frac{2U}{\pi}\int_0^\infty \sin\left(xy/\sqrt{2\kappa t}\right)\mathrm{e}^{-\frac{1}{2}y^2}\frac{1}{y}\,\mathrm{d}y \qquad \text{(putting } y^2 = 2rt\text{).} \end{aligned}$$

This can be rewritten as

$$u(x,t) = U\left[1 - \operatorname{erf}\left(\frac{x}{2\sqrt{\kappa t}}\right)\right],$$

where $\operatorname{erf}(x)$ is the **error function** defined by

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{e}^{-y^2}\,\mathrm{d}y.$$

To obtain this form, note that

$$\int_0^\infty \mathrm{e}^{-\frac{1}{2}y^2}\cos vy\,\mathrm{d}y = \sqrt{\pi/2}\mathrm{e}^{-\frac{1}{2}v^2}$$

(cf. Exercise 22.3 below) and integrate with respect to $v$ over $[0, x/\sqrt{2\kappa t}]$.

If the boundary condition $u(0,y) = U$ is replaced by the more general condition $u(0,t) = g(t)$, the Convolution theorem enables $u(x,t)$ to be obtained in terms of an integral (cf. Example 21.20).

### Tactical tip

⊙ We used the Laplace transform on the variable $t$. If we had operated by $\mathfrak{L}$ on $x$, we would have needed to know $u_x(0,t)$ (which we were not given) in order to apply 21.6. A (good) alternative approach to this example is via the sine transform (see 22.16 and Exercise 22.7).

## Exercises

**Exercises from the text.** Prove the claims in 21.5. Verify the entries in Table 21.1.

21.1 Find the Laplace transform of

$$\text{(i) } t(t^2-1), \qquad \text{(ii) } t(\cos t)e^{-t},$$
$$\text{(iii) } \cosh t \cos t, \qquad \text{(iv) } \chi_{[0,T]} \quad (T>0).$$

21.2 The $n$th Laguerre polynomial $L_n(t)$ is defined (for $n = 0, 1, \ldots$) by

$$L_n(t) = \frac{e^t}{n!}\left(\frac{d}{dt}\right)^n (t^n e^{-t}).$$

Prove that the Laplace transform of $L_n(t)$ is $(p-1)^n/p^{n+1}$. Deduce that

$$t\frac{d^2}{dt^2}L_n + (1-t)\frac{d}{dt}L_n + nL_n = 0.$$

21.3 Prove by induction that

$$\mathfrak{L}[(1-e^{\alpha t})^n] = \frac{n!\alpha^n}{p(p+\alpha)\ldots(p+n\alpha)} \qquad (n = 0, 1, 2, \ldots).$$

Hence compute $\mathfrak{L}[\sin^n t]$.

21.4 Assume that $f$ is defined on $[0,\infty)$ and is periodic of period $T$. Assuming the Laplace transform of $f$ exists, prove that

$$\overline{f}(p) = \frac{1}{1-pT}\int_0^T f(y)e^{-py}\,dy.$$

Now let

$$f(t) = \begin{cases} 2t/T & (0 \leqslant t \leqslant T/2), \\ 2(1-t/T) & (T/2 \leqslant t \leqslant T). \end{cases}$$

Show that, for $t \geqslant 0$,

$$f(t) = \frac{2}{T}\left(tH(t) + 2\sum_{n=1}^{\infty}(-1)^n(t - \tfrac{1}{2}nT)H(t - \tfrac{1}{2}nT)\right),$$

where $H$ is the Heaviside function (see 21.5(4)).

21.5 Verify that the Laplace inversion theorem gives the expected results for the functions listed in Table 21.1.

21.6 Find the inverse Laplace transform of

(i) $(p(p+1)(p+2))^{-1}$, (ii) $(p^2-1)^{-2}$,
(iii) $6(p^4+10p^2+9)^{-1}$, (iv) $((p^2+4)(p^2+1)^2)^{-1}$,
(v) $2p(p^4+1)^{-1}$, (vi) $2p^2(p^4+1)^{-1}$.

21.7 Let $f(t) = t^{-a}$, where $0 < a < 1$. For $\operatorname{Re} p > 0$, compute the Laplace transform $\overline{f}(p)$ in terms of the constant

$$C_a := \int_0^\infty u^{-a}\mathrm{e}^{-u}\,\mathrm{d}u.$$

Use 15.11 and 21.7 to obtain $\overline{f}(p)$ for $\operatorname{Re} p > 0$. Apply the inversion theorem to prove that

$$C_a C_{1-a} = \pi \operatorname{cosec} \pi a.$$

21.8 Solve the integral equation

$$y(t) = 1 + \int_0^t x\mathrm{e}^{-x}y(t-x)\,\mathrm{d}x \qquad (t > 0).$$

21.9 Use the Laplace transform to solve, for $t > 0$, the equation

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 6\frac{\mathrm{d}y}{\mathrm{d}t} + 13y = 0,$$

subject to $y(0) = y'(0) = 1$.

21.10 Use the Laplace transform to solve

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} - 5\frac{\mathrm{d}y}{\mathrm{d}t} + 4y = \begin{cases} 1 & \text{if } 0 < t \leqslant b, \\ -1 & \text{if } b < t \leqslant 2b, \\ 0 & \text{if } 2b < t, \end{cases}$$

with $y(0) = 0$, $y'(0) = 1$.

21.11 Solve, for $t \geqslant 0$, the simultaneous equations

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2n\frac{\mathrm{d}x}{\mathrm{d}t} + n^2x = 0, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} + 2n\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + n^2y = \mu\frac{\mathrm{d}x}{\mathrm{d}t},$$

where $\mu$ is a constant, $x(0) = x'(0) = 1$ and $y(0) = y'(0) = 0$.

21.12 Suppose that $x$, $y$, and $z$ are functions on $[0, \infty)$ such that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = bz - cy, \quad \frac{\mathrm{d}y}{\mathrm{d}t} = cx - az, \quad \frac{\mathrm{d}z}{\mathrm{d}t} = ay - bx,$$

where $a$, $b$, and $c$ are constants, and suppose that $x(0) = 1$ and $y(0) = z(0) = 0$. Show that

$$x(t) = \big(a^2 + (b^2 + c^2)\cos\omega t\big)/\omega^2, \quad \text{where } \omega^2 = a^2 + b^2 + c^2.$$

21.13 Use the Laplace transform to find a solution to the differential equation

$$tf''(t) + (1+t)f'(t) + f(t) = t^2 \quad (t \geqslant 0).$$

21.14 Suppose that $f(t)$ is a solution of the differential equation

$$f''(t) = \left(1 - \frac{4}{t} + \frac{2}{t^2}\right) f(t)$$

on $[0, \infty)$. Assuming that appropriate technical conditions are satisfied by $f$, show that the Laplace transform $\overline{f}(p)$ is a multiple of $(p+1)^{-3}$. Hence find a non-trivial solution to the differential equation.

21.15 Suppose that $f$ satisfies $f'(t) = f(kt)$ $(t > 0)$, where $0 < k < 1$, and $f(0) = 1$. Prove that

$$f(t) = \sum_{n=0}^{\infty} \frac{k^{n(n-1)/2}}{n!} t^n.$$

21.16 The functions $u_0, u_1, u_2, \ldots$ are related by the equations

$$u_n'(t) = u_{n-1}(t) - u_n(t) \quad (n \geqslant 1,\ t \geqslant 0).$$

Use the Laplace transform to prove that

$$u_n(t) = \int_0^t \varphi_{n-1}(t-x)u_0(x)\,\mathrm{d}x + \sum_{r=1}^{n} \varphi_{n-r}(t)u_r(0) \quad (n \geqslant 1,\ t \geqslant 0),$$

where the functions $\varphi_1, \varphi_2, \ldots$ are to be determined.

21.17 Use the Laplace transform to find $u(x,t)$ satisfying, for $x > 0$ and $t > 0$,

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = te^{-x},$$
$$u(x,0) = 0, \quad u_t(x,0) = x, \quad u(0,t) = 1 - e^{-t}.$$

21.18 Assume that $u(x,t)$ is defined in $\{ (x,t) \in \mathbb{R}^2 \,:\, x \geqslant 0,\, t \geqslant 0 \}$ and is a solution to the boundary value problem

$$x\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} = 1, \qquad u(1,t) = u(x,0) = 1.$$

By operating by the Laplace transform on the variable $t$, with $x$ fixed, prove that

$$\overline{u}(p,y) = \frac{1}{p^2} + \frac{1}{p} - \frac{e^{-pv}}{p^2}, \quad \text{where } x = e^v.$$

Deduce, with the aid of 21.5(4), that

$$u(x,t) = \begin{cases} 1 + t & \text{if } e^t < x, \\ 1 + \log x & \text{if } e^t \geqslant x. \end{cases}$$

21.19 The function $u(x,t)$ is continuous in $\{ (x,t) \,:\, x \geqslant 0,\, t \geqslant 0 \}$ and satisfies

(i) $u_{tt} = c^2 u_{xx} \quad (x > 0,\, t > 0)$;

(ii) $u(x,0) = u_t(x,0) = 0 \ (x > 0)$;

(iii) $\dfrac{d^2}{dt^2}u(0,t) + \mu^2 u(0,t) = \dfrac{2c^2}{b}u_x(0,t) \quad (\mu > c/b)$;

(iv) $u(0,0) = 0, \quad \left[\dfrac{d}{dt}u(0,t)\right]_{t=0} = U,$

where $c$, $\mu$, $b$, and $U$ are constants. Making such technical assumptions as you need, obtain $u(x,t)$, and verify that the solution you have found does satisfy the given conditions and any additional conditions imposed.

# 22 The Fourier transform

The Fourier transform is, from a theoretical point of view, more fundamental than the Laplace transform; we introduced the latter first because it has a greater wealth of elementary applications. Here we concentrate on the evaluation, by contour integration techniques, of Fourier transforms arising out of probability theory. We also present sufficient basic theory, parallel to that we presented for the Laplace transform, to be able to demonstrate how the Fourier transform can be used to solve certain ordinary and partial differential equations. A lively account of applications can be found in **[27]**. For the theory, see **[2]** or **[6]** (Lebesgue integral approach) or **[1]** (Riemann integral).

## Introducing the Fourier transform

**22.1 Definition (Fourier transform).** Let $f$ be a real- or complex-valued function on $\mathbb{R}$ such that $f \in \mathcal{I}(\mathbb{R})$; for the interpretation of $\mathcal{I}(\mathbb{R})$ see the technical note 21.1. The Fourier transform of $f$ is defined, for all real $s$, by

$$(\mathfrak{F}f)(s) = \widehat{f}(s) = \int_{-\infty}^{\infty} f(x)\mathrm{e}^{-\mathrm{i}sx}\,\mathrm{d}x.$$

Variants on this appear in some books: $\mathrm{e}^{\mathrm{i}sx}$ instead of $\mathrm{e}^{-\mathrm{i}sx}$ or an inserted normalization factor of $1/\sqrt{(2\pi)}$.

**22.2 Comparison of the Laplace and Fourier transforms.** The Laplace transform can be regarded as a special case of the Fourier transform, a fact we exploit in 22.9 when we derive its inversion theorem. To see the connection, write $p = u + \mathrm{i}s$ and suppose that $\overline{f}(p)$ exists. Let

$$g(x) = \mathrm{e}^{-ux} f(x)\chi_{[0,\infty)}(x).$$

Then

$$\widehat{g}(s) = \int_{0}^{\infty} \mathrm{e}^{-ux} f(x)\mathrm{e}^{-\mathrm{i}sx}\,\mathrm{d}x = \overline{f}(p).$$

In the definition of the Laplace transform, the factor $\mathrm{e}^{-px}$ has modulus $\mathrm{e}^{-\operatorname{Re} px}$; this is a negative exponential for $\operatorname{Re} p > 0$, and so decays fast as $x \to \infty$. For the Fourier transform, the exponential factor $\mathrm{e}^{-\mathrm{i}sx}$ has modulus 1 when $s$ is real, and so neither helps nor hinders the convergence of the integral defining $\widehat{f}$. Therefore, by comparison with the Laplace transform, we expect the Fourier transform to exist only for a relatively restricted class of functions.

The following result is entirely elementary.

**22.3 New Fourier transforms from old.** Provided all the transforms exist

(1) $\mathfrak{F}[af(x) + bg(x)] = a\mathfrak{F}[f(x)] + b\mathfrak{F}[g(x)]$ $(a, b \in \mathbb{C})$;
(2) $\mathfrak{F}[f(x/a)] = a\widehat{f}(sa)$ $(a > 0)$;
(3) $\mathfrak{F}[\mathrm{e}^{-\mathrm{i}xa} f(x)] = \widehat{f}(x + a)$ $(a \in \mathbb{R})$.

The remaining results in this section parallel those already given for the Laplace transform and form, with 22.3, the basis for the application of Fourier transforms to the solution of differential equations. The required formulae are straightforward to derive at a calculational level; the justifications are exercises in integration theory, of varying technical difficulty. For further details, see, for example, [**6**], Section 33.17.

**22.4 The Fourier transform of a derivative.** Under appropriate conditions on $f$,

$$\mathfrak{F}[f^{(n)}(x)] = (\mathrm{i}s)^n \widehat{f}(s).$$

This is derived by repeated integration by parts. Sufficient assumptions for this to be valid are:

(i) $f, f', \ldots, f^{(n)} \in \mathcal{I}(\mathbb{R})$, and
(ii) $f^{(n)}$ is continuous.

These conditions ensure that we can legitimately integrate by parts $n$ times. They are strong enough to force $f^{(k)}(x)\mathrm{e}^{-\mathrm{i}sx} \to 0$ as $|x| \to \infty$ for $k = 0, \ldots, n-1$; we need this in order to dispose of the integrated terms. In applications, decay conditions like this are often natural in the context of the problem being modelled.

**22.5 The derivative of a Fourier transform.** Under conditions on $f$ sufficient to justify differentiation under the integral sign $n$ times,

$$\mathfrak{F}[x^n f(x)] = \mathrm{i}^n \widehat{f}^{(n)}(s).$$

**22.6 Smoothness and decay.** There is an interpretation of the two preceding results which is worth noting. Let $f$ be an integrable function on $\mathbb{R}$. Observe that $|x^n| \to \infty$ as $x \to \infty$, and that this happens progressively faster as $n$ increases. Thus an assertion that $x^n f(x)$ is integrable on $\mathbb{R}$ tells us something about the rate of decay of $f(x)$ to zero as $|x| \to \infty$. Also, the number of derivatives a real-valued function $f$ possesses indicates the smoothness of the curve we get when we draw the graph of $f$. What 22.5 and 22.5 tell us is that the smoother $f$ is then the faster its Fourier transform deacys near $\pm\infty$, and vice versa.

## Evaluation and inversion

**22.7 Inversion: preliminary comments.** By contrast with with the good behaviour of the transforms of smooth and rapidly decaying functions, the transforms of integrable functions in general can behave quite badly. Consider, for example, $f(x) = \mathrm{e}^{-x}\chi_{[0,\infty)}(x)$. Then elementary integration gives

$$\widehat{f}(s) = \frac{1}{1+\mathrm{i}s} = \frac{1}{1+s^2} - \frac{\mathrm{i}s}{1+s^2}.$$

Considering the integral of the imaginary part of this function, we see that the PV-integral

$$\lim_{R\to\infty} \int_{-R}^{R} \frac{s}{1+s^2}\,\mathrm{d}s$$

exists but the improper integral $\lim_{R,S\to\infty} \int_{-S}^{R} s(1+s^2)^{-1}\,\mathrm{d}s$ does not. This suggests that a function $f$ may have better integrability properties than its transform $\widehat{f}$.

The inversion theorem for the Laplace transform gives a formula which allows a function to be recovered from its transform via a principal-value integral, in that case along some vertical line $\mathrm{Re}\,p = \sigma$ in $\mathbb{C}$. Since we can regard the Laplace transform as a special kind of Fourier transform, we would hope for a similar inversion formula for the latter.

The inversion theorem for the Fourier transform can be formulated in various ways, the main differences being in the class of functions considered. We present a version which is adequate for the kind of applications we wish to consider. A proof of this, in the setting of Lebesgue integration, can be found in **[6]**, Section 33.9. The argument given there shows how the inversion formula comes about, as well as paying attention to technical issues.

**22.8 Inversion theorem for the Fourier transform.** Let $f \in \mathcal{I}(\mathbb{R})$ and assume that $f$ is piecewise smooth (recall the definition in 21.9). Let

$$\widehat{f}(s) = \int_{-\infty}^{\infty} f(x)\mathrm{e}^{-\mathrm{i}sx}\,\mathrm{d}x.$$

Then

$$\frac{1}{2}\big(f(x+) + f(x-)\big) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{f}(s)\mathrm{e}^{\mathrm{i}sx}\,\mathrm{d}s.$$

Here the integral on the right-hand side may only exist in the weak sense of being a principal-value integral. If $f$ is continuous, the inversion formula takes the form

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{f}(s)\mathrm{e}^{\mathrm{i}sx}\,\mathrm{d}s.$$

The similarity of form of the integral formulae for $\widehat{f}(s)$ and $f(x)$ is please. As a consequence, evaluation and inversion of Fourier transforms are essentially equivalent processes.

**22.9 Deduction of the Laplace transform inversion theorem.** Assume that $f$ is piecewise smooth on $[0,\infty)$ and suppose that $\overline{f}(p) = \int_0^\infty f(t)\mathrm{e}^{-pt}\,\mathrm{d}t$ exists (in the sense defined in 21.2) for $\mathrm{Re}\,p > c \geqslant 0$, where $c$ is constant. Then, if $t > 0$,

$$\frac{1}{2}\big(f(t+) + f(t-)\big) = \frac{1}{2\pi\mathrm{i}}\lim_{R\to\infty}\int_{\alpha-\mathrm{i}R}^{\sigma+\mathrm{i}\rho} \overline{f}(p)\mathrm{e}^{pt}\,\mathrm{d}p \quad (\sigma > c).$$

To prove this, we take $p$ on the line of integration, so $p = \sigma + \mathrm{i}y$ and

$$\overline{f}(p) = \int_0^\infty \big(\mathrm{e}^{-\sigma t}f(t)\big)\mathrm{e}^{-\mathrm{i}yt}\,\mathrm{d}t.$$

We can apply the Fourier transform inversion theorem to $\mathrm{e}^{-\sigma t}f(t)\chi_{[0,\infty)}(t)$, which is piecewise smooth and belongs to $\mathcal{I}(\mathbb{R})$. This gives, for $\sigma > c$ and $t > 0$,

$$\frac{1}{2}\mathrm{e}^{-\sigma t}\big(f(t+) + f(t-)\big) = \frac{1}{2\pi}\lim_{R\to\infty}\int_{-R}^{R} \overline{f}(\sigma + \mathrm{i}y)\mathrm{e}^{\mathrm{i}yt}\,\mathrm{d}y.$$

We obtain the required inversion formula by replacing $\sigma + \mathrm{i}y$ by $p$. $\square$

Informally, the derivation of the following result is similar to that given earlier for the Laplace convolution (21.13). It involves interchanging the order in which repeated integrals are evaluated; the conditions stipulated are sufficient to validate this. It is part of the conclusion that the convolution integral is well defined.

**22.10 Convolution theorem for the Fourier transform.** Let $f, g \in \mathcal{I}(\mathbb{R})$. Then $\widehat{f}(s)\widehat{g}(s) = \widehat{h}(s)$ $(s \in \mathbb{R})$, where the **convolution** $h$ is defined by

$$h(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\,\mathrm{d}y.$$

## Applications

Our examples of computing and inverting Fourier transforms are drawn from probability theory. We consider some fundamental probability distributions on $\mathbb{R}$, and compute their characteristic functions. Such functions encode information about, for example, the moments of the associated distributions. For a probability distribution possessing a density function, $f$ say, the characteristic function is simply the Fourier transform $\widehat{f}$. Finding these transforms involves many of the techniques of contour integration presented in Chapter 20.

**22.11 Cauchy distribution.** Here we have $f(x) = 1/(\pi(1+x^2))$. The Fourier transform is

$$\widehat{f}(s) = \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\mathrm{i}sx}}{\pi(1+x^2)}\,\mathrm{d}x \quad (x \in \mathbb{R}).$$

We evaluate this integral by the method discussed for Example 20.5. When $z = R\mathrm{e}^{\mathrm{i}\theta}$, we have $\left|\mathrm{e}^{-\mathrm{i}sz}\right| = \mathrm{e}^{Rs\sin\theta}$. This leads us to use a semicircular contour, in the upper half-plane when $s \leqslant 0$ and in the lower half-plane when $s > 0$. This ensures that in either case the modulus of $\mathrm{e}^{-\mathrm{i}sx}$ on the semicircular arc is a negative exponential and so is bounded above by 1 (recall 19.4).

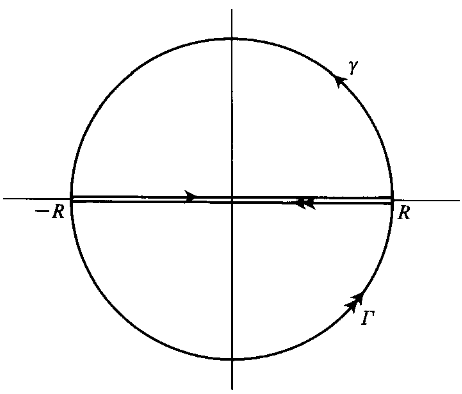



**Figure 22.1** Up and down

We arrive at

$$\widehat{f}(s) = \begin{cases} -2\pi \mathrm{i} \operatorname{res}\left\{ \dfrac{\mathrm{e}^{-\mathrm{i}sx}}{\pi(1+x^2)}; \mathrm{i} \right\} = \mathrm{e}^{s} & (s \leqslant 0), \\ 2\pi \mathrm{i} \operatorname{res}\left\{ \dfrac{\mathrm{e}^{-\mathrm{i}sx}}{\pi(1+x^2)}; -\mathrm{i} \right\} = \mathrm{e}^{-s} & (s > 0), \end{cases}$$

that is, $\widehat{f}(s) = \mathrm{e}^{-|s|}$ $(s \in \mathbb{R})$.

In this case the inversion integral

$$\frac{1}{2\pi} \lim_{R\to\infty} \int_{-R}^{R} \mathrm{e}^{-|s|}\mathrm{e}^{\mathrm{i}sx}\,\mathrm{d}x$$

can be directly computed. This we do by first splitting the range of integration into $(-\infty, 0]$ and $[0, \infty)$ and then evaluating the integrals of the real and imaginary parts of $\mathrm{e}^{-|s|}\mathrm{e}^{\mathrm{i}sx}$ on each of these intervals. We obtain $(\pi(1+x^2))^{-1}$, as the inversion theorem leads us to expect.

**22.12 Normal distribution.** Let

$$f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}.$$

We claim that $\widehat{f}(s) = \mathrm{e}^{\frac{1}{2}s^2}$. We have

$$\sqrt{2\pi}\widehat{f}(s) = \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}x^2}\mathrm{e}^{-\mathrm{i}sx}\,\mathrm{d}x = \mathrm{e}^{-\frac{1}{2}s^2} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}(x+\mathrm{i}s)^2}\,\mathrm{d}x.$$

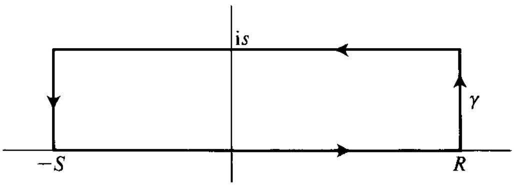



**Figure 22.2** Contour for the normal distribution integral

Integrate $\mathrm{e}^{-\frac{1}{2}z^2}$ round the rectangle with vertices at $-S$, $R$, $R + \mathrm{i}s$, $-S + \mathrm{i}s$ (shown in Fig. 22.2 in the case $s > 0$). By Cauchy's theorem,

$$\int_{-S}^{R} \mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x + \int_{0}^{s} \mathrm{e}^{-\frac{1}{2}(R+\mathrm{i}y)^2}\,\mathrm{i}\,\mathrm{d}y + \int_{R}^{-S} \mathrm{e}^{-\frac{1}{2}(x+\mathrm{i}s)^2}\,\mathrm{d}x + \int_{s}^{0} \mathrm{e}^{-\frac{1}{2}(-S+\mathrm{i}y)^2}\,\mathrm{i}\,\mathrm{d}y = 0.$$

Here

$$\left|\int_0^s e^{-\frac{1}{2}(R+iy)^2}\, i\, dy\right| \leqslant e^{-\frac{1}{2}R^2}\int_0^{|s|} e^{\frac{1}{2}y^2}\, dy \to 0 \text{ as } R\to\infty,$$

and similarly

$$\int_0^s e^{-\frac{1}{2}(-S+iy)^2}\, i\, dy \to 0 \text{ as } S\to\infty.$$

Hence, letting $R$ and $S$ tend to $\infty$,

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\, dx = \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x+is)^2}\, dx.$$

The left-hand side is $\sqrt{2\pi}$, by 20.15, so $\widehat{f}(s) = e^{-\frac{1}{2}s^2}$, as required.

This is an example where the answer appears to come out by making the complex substitution of $x + is$ for $x$. This procedure can be made respectable by contour integration; see 20.16.

Symmetry shows that the result obtained is consistent with the inversion theorem.

**22.13 Gamma distribution.** For $\lambda > 0$ and $t > 0$, let

$$f(x) = \frac{1}{\Gamma(t)}\lambda^t x^{t-1} e^{-\lambda x}\chi_{[0,\infty)}(x),$$

where

$$\Gamma(t) := \int_0^{\infty} x^{t-1} e^{-x}\, dx$$

defines the **gamma function.** We shall show that

$$\widehat{f}(s) = \left(\frac{\lambda}{\lambda + is}\right)^t,$$

where the right-hand side is a suitable branch of the power.

When $t$ is an integer we may integrate round a sector, motivated by the formal substitution of $(\lambda+is)x$ for $x$. In the general case we have to contend with a multifunction. We work in the plane cut along the negative real axis and take $z^{t-1} = \left|z^{t-1}\right| e^{i\theta(t-1)}$ for $z = |z|e^{i\theta}$ $(-\pi < \theta \leqslant \pi)$. Integrate $g(z) = z^{t-1}e^{-z}$ round the contour shown in Fig. 22.3. On CD, $z = (\lambda + is)u$ with $u > 0$, so

$$\int_{DC} g(z)\, dz = -(\lambda + is)^t \int_{\varepsilon}^{R} u^{t-1} e^{-(\lambda+is)u}\, du,$$

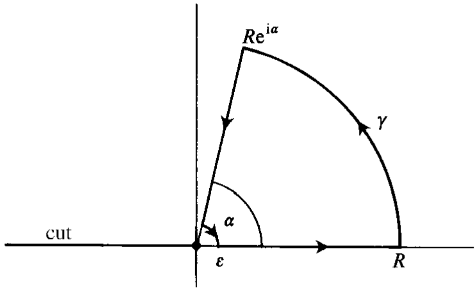



**Figure 22.3** Contour for the gamma distribution integral

while

$$\int_{AB} g(z)\,\mathrm{d}z = \int_{\varepsilon}^{R} x^{t-1}\mathrm{e}^{-x}\,\mathrm{d}x.$$

Also

$$\begin{aligned}\left|\int_{BC} g(z)\,\mathrm{d}z\right| &\leqslant \int_0^{|\alpha|} \left|(R\mathrm{e}^{\mathrm{i}\theta})^{t-1}\mathrm{e}^{-R\mathrm{e}^{\mathrm{i}\theta}} R\mathrm{i}\mathrm{e}^{\mathrm{i}\theta}\right| \mathrm{d}\theta \quad (\text{where } \tan\alpha = s/\lambda)\\ &\leqslant \int_0^{|\alpha|} R^t\mathrm{e}^{-R\cos\theta}\,\mathrm{d}\theta\\ &\leqslant |\alpha|\,R^t\mathrm{e}^{-R\cos\alpha},\end{aligned}$$

which tends to zero as $R \to \infty$, and

$$\left|\int_{DA} g(z)\,\mathrm{d}z\right| \leqslant |\alpha|\,\varepsilon^t\mathrm{e}^{-\varepsilon\cos\alpha},$$

which tends to zero as $\varepsilon \to 0$, since $t > 0$.

Apply Cauchy's theorem and take the limit as $R \to \infty$ and $\varepsilon \to 0$ to obtain

$$(\lambda + \mathrm{i}s)^t \int_0^\infty u^{t-1}\mathrm{e}^{-(\lambda+\mathrm{i}s)u}\,\mathrm{d}u = \int_0^\infty x^{t-1}\mathrm{e}^{-x}\,\mathrm{d}x.$$

Therefore

$$\widehat{f}(s) = \left(\frac{\lambda}{\lambda + \mathrm{i}s}\right)^t := \left|\frac{\lambda}{\lambda+\mathrm{i}s}\right|^t \mathrm{e}^{\mathrm{i}t\alpha},$$

where $\tan\alpha = s/\lambda$, $s = |s|\,\mathrm{e}^{\mathrm{i}\theta}$, and $-\frac{1}{2}\pi < \alpha < \frac{1}{2}\pi$.

**22.14 Example (an ordinary differential equation).** Consider

$$f''(x) - f(x) = e^{-x^2} \quad (x \in \mathbb{R}),$$

where we assume that $f(x), f'(x), f''(x)$ all belong to $\mathcal{I}(\mathbb{R})$ and tend to zero as $|x| \to \infty$. Operate by $\mathfrak{F}$. Using results in 22.3 and 22.4, we obtain

$$-s^2\widehat{f}(s) - \widehat{f}(s) = \mathfrak{F}[e^{-x^2}].$$

Hence, by 22.11, 22.10, and 22.8,

$$f(x) = -\pi \int_{-\infty}^{\infty} e^{-|y-x|-y^2}\, dy.$$

(Note that we could have calculated the Fourier transform of $e^{-x^2}$ using 22.3(2) and 22.12, but did not need to do so.)

**22.15 Example (a boundary value problem: Laplace's equation in a half-plane).** Assume $u(x,y)$ is defined and continuous on $\{ (x,y) \in \mathbb{R}^2 : y \geqslant 0 \}$ and that it satisfies

(i) $u_{xx} + u_{yy} = 0$;

(ii) $u(x,0) = f(x)$ $(x \in \mathbb{R})$, where $f$ is integrable on $\mathbb{R}$.

We shall solve this partial differential equation for $u$, subject to suitable restrictions on the behaviour of $u(x,y)$ for large values of $r = (x^2+y^2)^{1/2}$.

We operate by the Fourier transform on the variable $x$ and write, for fixed $y$,

$$\widehat{u}(s,y) = \int_{-\infty}^{\infty} u(x,y)e^{-isx}\, dx.$$

The partial differential equation (i) transforms into

$$\frac{d^2\widehat{u}}{dy^2} = s^2\widehat{u}.$$

In deriving this we have used 22.4 and have assumed that the derivatives of $\widehat{u}$ with respect to $y$ can be obtained by differentiation under the integral sign. The boundary condition (ii) transforms to $\widehat{u}(s,0) = \widehat{f}(s)$ and, provided $u(x,y)$ decays sufficiently rapidly as $r \to \infty$, we have $\widehat{u}(s,y) \to 0$ as $y \to \infty$, for each $s$. Then

$$\widehat{u}(s,y) = \widehat{f}(s)e^{-|s|y}.$$

By the Convolution theorem and Example 22.11,

$$u(x,y) = \frac{y}{\pi}\int_{-\infty}^{\infty} \frac{f(t)}{(x-t)^2+y^2}\,\mathrm{d}t.$$

We have here the Poisson integral for a half-plane. This is considered further in 23.16.

**22.16 Postscript: sine and cosine transforms.** As books on Fourier analysis explain, the Fourier transform can be regarded as the limiting case of a Fourier series when the periodicty of the function tends to zero. There are therefore parallels between the theory of Fourier series and that of Fourier transforms.

Consider the Fourier series of a $2\pi$-periodic function $f$ on $\mathbb{R}$. If $f$ is even (odd), then the Fourier series contains only cosine terms (only sine terms). In a similar way, the exponential factor in the Fourier transform of an even (odd) function is replaced by a cosine (sine) term. Since any function on $[0,\infty)$ can be extended to become an even or an odd function on $\mathbb{R}$, sine and cosine transforms provide an alternative to the Laplace transform. They can be used to solve certain ordinary and partial differential equations. Which transform is appropriate depends on the form of the initial or boundary conditions. See Exercises 22.3 and 22.7.

## Exercises

**Exercises from the text.** Verify the formulae given in 22.3, 22.4, 22.5, and 22.10 [giving justifications if you have the technical knowledge to do so].

22.1 Compute the Fourier transform of $f$, where

$$f(x) = \left(1 - \frac{|x|}{\lambda}\right)\chi_{[-\lambda,\lambda]}(x).$$

Hence find the value of

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2}\,\mathrm{d}x.$$

22.2 Find the inverse Fourier transform of

(i) $(1+\mathrm{i}s)^{-1}$, (ii) $(1-s^2)(1+s^2)^{-2}$,

(iii) $s\mathrm{e}^{-s}\chi_{[0,\infty)} - s\mathrm{e}^{s}\chi_{(-\infty,0)}$; (iv) $s^{-1}\sin s$.

22.3 Suppose that $f$ satisfies the hypotheses of the Fourier inversion theorem (22.8) and that $f(x) = f(-x)$ for all $x$. Show that

$$\frac{1}{2}\big(f(x+) + f(x-)\big) = \frac{2}{\pi}\int_0^\infty \cos vx \left(\int_0^\infty f(y)\cos vy \, dy\right) dv.$$

Hence evaluate, for $a \in \mathbb{R}$,

$$\text{(i)} \int_0^\infty e^{-v^2} \cos 2av \, dv, \qquad \text{(ii)} \int_0^\infty \frac{\sin av \cos av}{v} \, dv.$$

22.4 Use the result of 22.12 to show that

$$\int_0^\infty e^{-\alpha u^2} \cos \beta u \, du = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}} e^{-\beta^2/(4\alpha)} \quad (\alpha > 0).$$

22.5 Use the Fourier transform to find a solution to the differential equation

$$f''(x) - 2f(x) = e^{-|x|} \qquad (x \in \mathbb{R}),$$

using (i) partial fractions and inspection, (ii) the Fourier inversion theorem and contour integration, and (iii) the Convolution theorem, to obtain $f$ from $\widehat{f}$. (The primary purpose of this exercise is to give practice in inversion techniques; the given differential equation can be solved by more elementary means.)

22.6 Use the Fourier transform to solve

$$f''(x) + 2f'(x) + f(x) = g(x) \quad (x \in \mathbb{R}),$$

expressing the solution as a convolution.

22.7 Let $f$ be an integrable function on $[0, \infty)$ and extend $f$ to an integrable function $F$ on $\mathbb{R}$ by defining

$$F(x) = \begin{cases} f(x) & (x \geqslant 0), \\ -f(-x) & (x < 0). \end{cases}$$

By applying the Fourier inversion theorem to $F$, show that, if $f$ is suitably smooth, $f(t)$ can be expressed in terms of its sine transform

$$\int_0^\infty f(t) \sin ts \, dt.$$

Use the sine transform to give an alternative solution to Example 21.22.

# 23 Harmonic functions and conformal mapping

This chapter first explores the connections between the theory of holomorphic functions and that of harmonic functions. The latter are functions which arise, locally, as the real parts of holomorphic functions. Because harmonic functions satisfy Laplace's equation in two dimensions, they occur widely in applied mathematics. We cannot in the space available give more than a few hints as to why this is so. We aim principally to provide links to applications-oriented texts, and in particular to show that techniques from complex analysis are of relevance. One such technique is conformal mapping and we conclude the chapter with a brief discussion of some mappings which arise out of problems in fluid dynamics.

Readers interested primarily in applications may wish to skip over the proofs of the theoretical results.

## Harmonic functions

Laplace's equation $u_{xx} + u_{yy} = 0$ is of fundamental importance in the mathematical modelling of 2-dimensional physical problems concerning fluid flow, steady heat conduction, electrostatics, and other phenomena. In the context of fluid flow it is assumed that the fluid is incompressible and inviscid and that the flow is steady and irrotational.

Harmonic functions are formally defined in 23.4. Before introducing them we make some brief and informal remarks about the modelling of fluid flow.

**23.1 Complex potential.** As we mentioned in Chapter 5, and shall prove in 23.3, we have, for a holomorphic function $f = u + iv$,

$$u_{xx} + u_{yy} = 0 = v_{xx} + v_{yy},$$

so that $u$ and $v$ satisfy Laplace's equation. In the other direction, suppose we have a solution $\phi = \phi(x, y)$ of Laplace's equation in two dimensions and suppose that $\phi$ is expressible as $\operatorname{Re} w$ for some holomorphic function $w$. Therefore $\phi$ possesses a **harmonic conjugate** $\psi = \operatorname{Im} w$; this also satisfies Laplace's equation.

In the context of 2-dimensional fluid flow, $\phi$ is the **velocity potential** and $\psi$ the **stream function.** The function $w = \phi + \mathrm{i}\psi$ is then called the **complex potential.** Its derivative

$$\mathbf{V}(z) := w'(z) = \phi_x(x, y) + \mathrm{i}\psi_x(x, y)$$

determines a vector field which models the velocity of fluid motion. Elementary vector calculus shows that each curve $\phi =$ constant is orthogonal to each curve $\psi =$ constant. The latter curves follow the direction of the fluid flow and are known as **streamlines.**

A basic problem is to solve $\phi_{xx} + \phi_{yy} = 0$ in some region $G$ in $\mathbb{R}^2$, where $u$ is to be continuous on $\overline{G}$ and certain conditions involving the normal derivative $\partial\phi/\partial n$ are to be satisfied on the boundary $\partial G$. We might, for example, wish to consider fluid flow through a cylindrical pipe whose cross-section is given by a simple closed curve $C$ in $\mathbb{R}^2$, with $G$ the region bounded by $C$, or the flow past a cylindrical obstacle with $C$ as its cross-sectional boundary. In either case, $C$ has to be a streamline.

Working with a complex potential is simpler than treating the velocity potential and stream function separately, and allows us to draw on the theory of complex functions.

- The key theorems of complex analysis relate closely to theorems concerning harmonic functions, the most fundamental relationship being that between Cauchy's theorem and Green's theorem (see 23.11).
- Invertible conformal mappings can be used to transform geometrically complex configurations to simpler ones, and back again. Crucially, conformal maps preserve harmonicity, so that the solution of simple fluid flow problems can lead to the solution of more complicated ones.
- In simple cases of flow past an obstacle, the complex velocity $\mathbf{V}(z)$ and complex potential $w(z) = \phi + \mathrm{i}\psi$ can be found explicitly by taking a Laurent expansion of $\mathbf{V}$ and matching this to conditions required to hold on the boundary of the obstacle and at infinity.
- Quantities relating to the flow, such as forces, can be expressed in terms of the complex potential, and techniques of complex analysis employed to find these.

To illustrate the last point, we introduce Blasius's theorem. For steady flow in the absence of gravity, Bernoulli's equation,

$$p(z) + \frac{1}{2}\rho\,|\mathbf{V}(z)|^2 = p_0,$$

relates the fluid pressure $p(z)$ and velocity $\mathbf{V}(z)$; here $\rho$ is the density of the fluid (assumed constant) and $p_0$ is a constant. This leads to Blasius's theorem. For a fixed body which is described parametrically by a positively oriented contour [simple closed path] $\gamma$. The theorem gives $F = (X, Y)$, the force per unit length, and $M$, the moment (or torque), on the body:

$$X - \mathrm{i}Y = \frac{1}{2}\mathrm{i}\rho \int_\gamma \left(\frac{\mathrm{d}w}{\mathrm{d}z}\right)^2 \mathrm{d}z \quad \text{and} \quad M = \frac{1}{2}\rho \operatorname{Re} \int_\gamma z \left(\frac{\mathrm{d}w}{\mathrm{d}z}\right)^2 \mathrm{d}z;$$

$X$ can be interpreted as the drag on the body and $Y$ as the lift. Exercise 23.5 seeks $F$ and $M$ for some particular complex potentials.

**23.2 Examples (complex potential).**

- $w(z) = Uz$ is the complex potential for a uniform flow of speed $U$ parallel to the real axis.
- Let $w(z) = U\left(z + \dfrac{a^2}{z}\right)$ $(|z| > a)$. This is certainly holomorphic, so that its real and imaginary parts are harmonic. These are given in polar coordinates by

$$\phi(r, \theta) = U\left(r + \frac{a^2}{r}\right)\cos\theta \quad \text{and} \quad \psi(r, \theta) = U\left(r - \frac{a^2}{r}\right)\sin\theta.$$

We have $\psi = 0$ for $\theta = 0$ and $\theta = \pi$, as well as on the circle $r = a$. For very large $r$, the streamlines approximate to straight lines parallel to the real axis. This example models uniform 2-dimensional fluid flow parallel to the $x$-axis, into which a circular cylinder of radius $a$ has been placed.

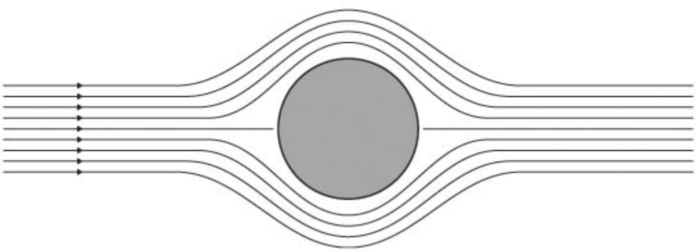


**Figure 23.1** Flow past a cylinder

In general, the **Milne–Thomson circle theorem** asserts that if a circular cylinder $|z| = a$ is inserted into a flow having complex potential $w$ then the new flow has complex potential given by $w(z) + \overline{w(a/\overline{z})}$.

We now present the theory of harmonic functions that we shall need.

**23.3 Holomorphy and the Cauchy–Riemann equations: a re-assessment.** We proved in 5.3 that if $f = u + \mathrm{i}v$ in an open set $G$ then $f' = u_x + \mathrm{i}v_x = -\mathrm{i}u_y + v_y$. These equations led us to the Cauchy–Riemann equations $u_x = v_y$ and $u_y = -v_x$. In 13.7 we proved that if $f \in \mathrm{H}(G)$ then in fact $f$ is infinitely differentiable in $G$. This implies that $u$ and $v$ have partial derivatives of all orders. As a consequence we have in particular:

- $u_x$, $u_y$, $v_x$, and $v_y$ are continuous, and
- the second-order partial derivatives of $u$ and $v$ exist and are continuous.

In 5.6 we proved that if the Cauchy–Riemann equations hold in an open set $G$ then $f := u + \mathrm{i}v$ is holomorphic in $G$, *so long as $u_x$, $u_y$, $v_x$, and $v_y$ are continuous.* We now see that this condition is in fact a necessary one. Furthermore, continuity of the second-order partial derivatives implies that $u_{xy} = u_{yx}$ and $v_{xy} = v_{yx}$. Hence

$$u_{xx} = v_{yx} = v_{xy} = -u_{xx} \quad \text{and} \quad v_{xx} = -u_{yx} = -u_{xy} = -v_{yy}.$$

**23.4 Definition (harmonic function).** Let $G$ be an open subset of $\mathbb{C}$ and identify $z = x + \mathrm{i}y \in \mathbb{C}$ with $(x, y) \in \mathbb{R}^2$. A function $u \colon G \to \mathbb{R}$ is **harmonic** in $G$ if

(i) $u$ has continuous second-order partial derivatives in $G$, and

(ii) $u$ satisfies Laplace's equation $u_{xx} + u_{yy} = 0$ in $G$.

We denote by $\mathcal{H}(G)$ the set of functions harmonic in $G$.

The facts in 23.3 are sufficiently important to be recorded as a theorem.

**23.5 Theorem (holomorphy and harmonicity).** Let $f$ be holomorphic in an open set $G$ and let

$$f(z) = u(x, y) + \mathrm{i}v(x, y) \quad (z = x + \mathrm{i}y \in G),$$

where $u$ and $v$ are real-valued. Then $u$ and $v$ are harmonic in $G$.

**23.6 Theorem (existence of a harmonic conjugate).** Let $G$ be an open disc [a simply connected region] and suppose that $u \in \mathcal{H}(G)$. Then there exists $v \in \mathcal{H}(G)$ such that $f = u + \mathrm{i}v \in \mathrm{H}(G)$, so $u$ is the real part of a holomorphic function in $G$.

**Proof** If $f$ exists then we must have $f'(z) = u_x + \mathrm{i}v_x = u_x - \mathrm{i}u_y$. Let $g(z) := u_x - \mathrm{i}u_y$. We shall apply Lemma 5.6 to $g$ to show that $g \in \mathrm{H}(G)$. The Antiderivative theorem (11.4 [12.6]), provides $F \in \mathrm{H}(G)$ with $F' = g$. Then

$$(u - \operatorname{Re} F)_x = (u - \operatorname{Re} F)_y = 0$$

in $G$, whence $u - \operatorname{Re} F$ is a real constant $k$ (see 5.12). Now let $f = F + k$. Then $f \in \mathrm{H}(G)$ and $\operatorname{Re} f = u$. □

**23.7 Finding a harmonic conjugate.** In practice, given $G$ and $u$, it is often possible to recognize a harmonic conjugate for $u$ at sight. Suppose, for example, that $u(x, y) = x - xy$. This is harmonic in $G = \mathbb{C}$. Then $u_x - \mathrm{i}u_y = 1 - y + \mathrm{i}x = 1 + \mathrm{i}z$. Hence $u = \operatorname{Re} f$, where $f(z) = z + \frac{1}{2}z^2$, and $v = y + xy$, is a harmonic conjugate for $u$.

Where we cannot spot how to choose $f$ we use the relation

$$f(w) - f(a) = \int_{\gamma(w)} f'(z)\,\mathrm{d}z = \int_{\gamma(w)} (u_x - \mathrm{i}u_y)\,\mathrm{d}z,$$

which holds for any polygonal path $\gamma(w)$ in $G$ joining a fixed point $a \in G$ to $w$. The Deformation theorem 11.9 [12.4] implies that the integrals are independent of the choice of path. In practice, it is often convenient to take a path consisting of horizontal and vertical line segments.

**23.8 Poisson integral formula (for holomorphic and for harmonic functions).**

(1) Let $f$ be holomorphic in an open disc containing $\overline{\mathrm{D}}(0; R)$. Then

$$f(r\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{(R^2 - r^2)}{(R^2 - 2Rr\cos(\theta - t) + r^2)} f(R\mathrm{e}^{\mathrm{i}t})\,\mathrm{d}t \quad (0 \leqslant r < R).$$

(2) Let $u$ be harmonic in an open disc $G$ containing $\overline{\mathrm{D}}(0; R)$. Then

$$u(r\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{(R^2 - r^2)}{(R^2 - 2Rr\cos(\theta - t) + r^2)} u(R\mathrm{e}^{R\mathrm{i}t})\,\mathrm{d}t \quad (0 \leqslant r < R).$$

**Proof** (1) Fix $z = re^{i\theta}$ and apply Cauchy's integral formula (13.1) to the product $fg$, where $g(w) = (R^2 - r^2)/(R^2 - w\bar{z})$. Then

$$f(z) = f(z)g(z) = \frac{(R^2 - r^2)}{2\pi i} \int_{\gamma(0;R)} \frac{f(w)}{(w - z)(R^2 - w\bar{z})} \, dw$$
$$= \frac{(R^2 - r^2)}{2\pi i} \int_0^{2\pi} \frac{f(Re^{it})}{(Re^{it} - re^{i\theta})(R^2 - Rre^{i(t-\theta)})} Rie^{it} \, dt.$$

With the aid of the identity $2\cos\alpha = e^{i\alpha} + e^{-i\alpha}$, this simplifies to give the integral required.

(2) Theorem 23.6 allows us to choose $f \in H(G)$ such that $u = \operatorname{Re} f$. The required formula is obtained from (1), the Poisson integral formula for $f$ by equating real parts. □

**23.9 Mean value property for harmonic functions.** Under the same hypotheses as in 23.8(2),

$$u(0) = \frac{1}{2\pi} \int_0^{2\pi} u(Re^{it}) \, dt.$$

**Proof** The formula is clearly a special case of that in 23.8(2). □

**Technical note** The hypotheses in 23.8 and 23.9 can be weakened. It is enough to assume that $u$ is harmonic in $D(0;1)$ and continuous on $\overline{D}(0;1)$. We can apply the preceding theorems to $u_\rho$, where $u_\rho(z) := u(\rho z)$ $(\rho < 1)$ and take the limit as $\rho$ increases to 1. [Full justification requires the fact that $u$ is uniformly continuous.]

**23.10 Maximum principle for harmonic functions.** Let $G$ be a bounded region and let $u$ be harmonic in $G$ and continuous on $\overline{G}$. Suppose that $u \leqslant M$ on $\partial G := \overline{G} \setminus G$, where $M$ is a constant. Then $u \leqslant M$ on $\overline{G}$, that is, $u$ attains its maximum on the boundary $\partial G$ of $G$.

**Outline proof** The Maximum principle is the harmonic counterpart of the Maximum modulus theorem (16.2), and can be proved in an analogous way. In brief, a local version is first derived from the Mean value property (cf. 16.1); then we show that $\{ z \in G : u(z) = M \}$ is either the whole of $G$ or the empty set, in the same manner as in the proof of the Identity theorem (15.8). □

**23.11 The relationship of Cauchy's theorem to Green's theorem.** Many readers will have seen Laplace's equation before, either in the context of mathematical modelling or of partial differential equations, and such results as the Mean value property and the Maximum principle may already be familiar. Such results—not restricted to the 2-dimensional setting—are usually derived using the apparatus of vector calculus. The proofs have their roots in Green's theorem (or, more fundamentally, in Stokes' theorem on differential forms). Green's theorem is a deep result. It is customarily presented without proof in introductory courses on vector calculus and is widely used in applied mathematics.

Let $f$ be holomorphic inside and on a closed path $\gamma$. Blithely ignoring any technical hurdles, put $\mathrm{d}z = \mathrm{d}x + \mathrm{i}\,\mathrm{d}y$ and so write $\int_\gamma f(z)\,\mathrm{d}z$ in terms of line integrals as

$$\int_\gamma (u\,\mathrm{d}x - v\,\mathrm{d}y) + \mathrm{i}\int_\gamma (u\,\mathrm{d}y + v\,\mathrm{d}x).$$

We then use Green's theorem to rewrite this as

$$\iint_{\mathrm{I}(\gamma)} (-v_x - v_y)\,\mathrm{d}x\,\mathrm{d}y + \mathrm{i}\iint_{\mathrm{I}(\gamma)} (u_x - v_y)\,\mathrm{d}x\,\mathrm{d}y.$$

The Cauchy–Riemann equations imply that both of these integrals are zero. So we conclude that $\int_\gamma f(z)\,\mathrm{d}z = 0$. Sadly, this is not quite the short cut to Cauchy's theorem that it might seem. To justify it, one must assume, or somehow prove otherwise, that $f'$ is continuous. But we deduced this from a consequence of Cauchy's theorem. The approach is, however, certainly of historical interest: it was the one used by Cauchy to derive the theorem that bears his name.

**23.12 The role of conformal mapping.** The philosophy behind the integral transform methods of the two preceding chapters is that of converting a given problem to one which is easier to solve, and then 'inverting' to solve the original problem. The same philosophy underlies the use of (invertible) conformal mappings in problems involving harmonic functions. Suppose that we have a region $G$ and a one-to-one conformal mapping $g$ of $G$ onto a simpler region, say $\mathrm{D}(0;1)$. We do not need to assume also that $g^{-1}$ is conformal since this comes free, by the Inverse function theorem (16.7). Significantly, the correspondence between $G$ and $\mathrm{D}(0;1)$ set up by $g$ and $g^{-1}$ goes beyond pure geometry. Because the maps are conformal, it turns out that they transfer harmonicity backwards and forwards too. Hence a boundary value problem for $G$ is converted to an equivalent boundary value problem for $\mathrm{D}(0;1)$, Streamlines for a fluid flow problem for an obstacle $\overline{G}$ correspond to streamlines for flow past a circular cylinder, and so on. The key to all this lies in an elementary lemma.

**23.13 Transfer lemma (composing a harmonic function with a holomorphic map).** Suppose that $G$ and $\widehat{G}$ are open sets, that $g\colon G \to \widehat{G}$ is holomorphic, and that $\widehat{u} \in \mathcal{H}(\widehat{G})$. Then $u\colon = \widehat{u} \circ g \in \mathcal{H}(G)$.

**Proof** We put $\xi + \mathrm{i}\eta = g(z) = g(x + \mathrm{i}y)$, so that $\widehat{u}(\xi, \eta) = u(x, y)$. Then $g'(z) = \xi_x + \mathrm{i}\eta_x = -\mathrm{i}\xi_y + \eta_y$. Straightforward partial differentiation shows that

$$u_{xx} + u_{yy} = |g'(z)|^2 \left(\widehat{u}_{\xi\xi} + \widehat{u}_{\eta\eta}\right) = 0. \quad \square$$

Note too that the second-order partial derivatives of $u$ with respect to $x$ and $y$ are continuous because $g'$ is continuous and $\widehat{u}$ has continuous second-order partial derivatives. $\square$

## The Dirichlet problem and its solution by conformal mapping

Let $G$ be a region. Suppose that we are given a real-valued continuous function $U$ on the boundary $\partial G = \overline{G} \setminus G$. Can we find a function $u$ such that $u$ is continuous on $\overline{G}$ and harmonic in $G$ and is such that $u = U$ on $\partial G$? This boundary value problem is known as the **Dirichlet problem.** The solution, if it exists, is unique: if $u_1$ and $u_2$ are both solutions, apply the Maximum principle to $u_1 - u_2$ and to $u_2 - u_1$ to prove that $u_1 \equiv u_2$ on $\overline{G}$.

The simpler the geometric configuration, the simpler, presumably, the Dirichlet problem will be. We now solve the problem for the simplest case of all: $G = \mathrm{D}(0; 1)$.

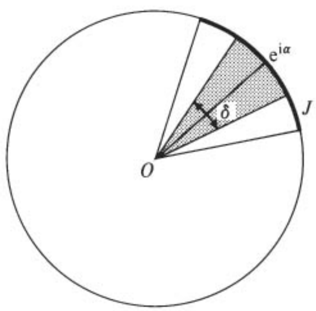



**Figure 23.2** Dirichlet problem for $\mathrm{D}(0; 1)$

**23.14 The Dirichlet problem for the unit disc.** Suppose that $U$ is a real-valued continuous function on the unit circle. Let

$$v(re^{i\theta}) := \frac{1}{2\pi}\int_0^{2\pi} \frac{(1-r^2)}{(1-2r\cos(\theta-t)+r^2)}U(e^{it})\,dt \quad (re^{i\theta} \in D(0;1))$$

be the **Poisson integral** of $U$. Define $u$ by

$$u(re^{i\theta}) := \begin{cases} v(re^{i\theta}) & (0 \leqslant r < 1), \\ U(e^{i\theta}) & (r = 1). \end{cases}$$

Then $u \in \mathcal{H}(D(0;1))$ and $u$ is continuous on $\overline{D}(0;1)$.

**Outline proof** The **Poisson kernel**

$$P_r(t) := \frac{1-r^2}{1-2r\cos(\theta-t)+r^2}$$

may be alternatively written as $\mathrm{Re}\big((w+z)/(w-z)\big)$ or as $(1-|z|^2)/\,|w-z|^2$, where $w = e^{it}$ and $z = re^{i\theta}$; see Exercise 1.13. For $|z| < 1$,

$$\begin{aligned} u(z) &= \mathrm{Re}\left(\frac{1}{2\pi i}\int_{\gamma(0;1)} \frac{w+z}{w(w-z)}U(w)\,dw\right) \\ &= \mathrm{Re}\left(\frac{1}{\pi i}\int_{\gamma(0;1)} \frac{U(w)}{w-z}dw\right) - \mathrm{Re}\left(\frac{1}{2\pi i}\int_{\gamma(0;1)} \frac{U(w)}{w}\,dw\right). \end{aligned}$$

The second term is constant, while the first is the real part of the derivative of a Cauchy integral, which is holomorphic (see 13.7). Hence $u$ is harmonic in $D(0;1)$.

To prove the continuity assertion we have to show that $v(re^{i\theta}) \to U(e^{i\alpha})$ as $re^{i\theta} \to e^{i\alpha}$. Putting $u$ equal to the constant function 1 in the Poisson integral formula gives

$$1 = \frac{1}{2\pi}\int_0^{2\pi} P_r(\theta-t)\,dt.$$

Hence

$$v(re^{i\theta}) - U(e^{i\alpha}) = \frac{1}{2\pi}\int_0^{2\pi} P_r(\theta-t)\big(U(e^{it}) - U(e^{i\alpha})\big)\,dt.$$

Let $\varepsilon > 0$. Continuity of $U$ implies that there exists $\delta > 0$ such that $\big|U(e^{it}) - U(e^{i\alpha})\big| < \varepsilon$ for all $t$ such that $e^{it}$ lies on the arc $J$ of the unit circle $T$ which contains $e^{i\alpha}$ and joins the points $e^{i(\alpha\pm\delta)}$. Let $K$ be the arc of $T$ complementary to $J$. Then for some $m > 0$ we have $|w-z| \geqslant m$ whenever $w \in K$ and $z$ lies in the shaded sector shown in Fig. 23.2.

Noting that $P_r(\theta - t) \geqslant 0$, we now have

$$\left|v(re^{i\theta}) - U(e^{i\alpha})\right| \leqslant \frac{1}{2\pi}\int_0^{2\pi} P_r(\theta - t)\left|U(e^{it}) - U(e^{i\alpha})\right| dt.$$

Elementary estimates show that the contribution from $J$ to this integral is less than $\varepsilon$, while, for $z$ in the shaded sector, the contribution from $K$ is bounded by a multiple of $(1-r^2)/m^2$ (remember that $U$ is bounded on the circle $T$). Hence we can make $\left|v(re^{i\theta}) - U(e^{i\alpha})\right| < 2\varepsilon$ by taking $\left|re^{i\theta} - e^{i\alpha}\right|$ small enough. □

The preceding result can easily be adapted to solve the Dirichlet problem in an arbitrary disc $D(a; R)$, by translating and rescaling. Transfer to other regions can be accomplished using conformal mapping.

**23.15 Solving the Dirichlet problem by conformal mapping.** Let $G$ be a region and suppose that we can find a one-to-one continuous map $g : \overline{G} \to \overline{D}(0;1)$ which maps $G$ conformally onto $D(0;1)$ and maps the boundary $\partial G$ onto the unit circle. [Some remarks on whether such a mapping $g$ exists in general can be found in 16.10.] Then we can use 23.14 and the Transfer lemma (23.13) to solve the Dirichlet problem for $G$: since $\widehat{U} := U \circ g^{-1} : T \to \mathbb{R}$ is a continuous function on the unit circle we can find a function $\widehat{u}$ on $\overline{D}(0;1)$ which is harmonic on $D(0;1)$ and which agrees with $\widehat{U}$ on $T$. Finally, $u := \widehat{u} \circ g$ is continuous on $\overline{G}$ and harmonic in $G$.

In practical problems, the boundary function $U$ is often piecewise continuous rather than continuous. Theorem 23.14 can be extended to cover this case.

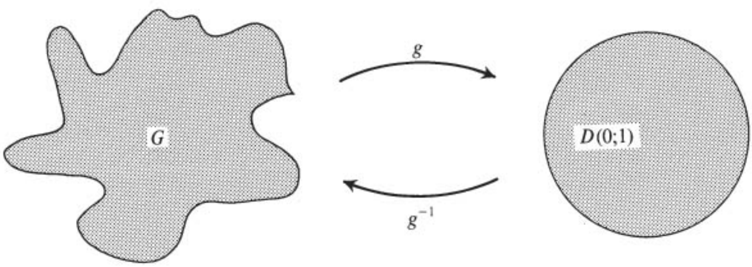



**Figure 23.3** The Dirichlet problem by conformal mapping

**23.16 The Dirichlet problem for a half-plane.** Let $G$ be the open upper half-plane $\Pi^+$ and suppose that $U$ is a real-valued continuous function on the real axis. Fix $z = x + \mathrm{i}y$ $(y > 0)$. Let $g(\zeta) := (\zeta - z)/(\zeta - \overline{z})$ $(\operatorname{Im}\zeta \geqslant 0)$. Then $g$ maps $G$ conformally one-to-one onto $\mathrm{D}(0;1)$. As in 23.15, let $u = \widehat{u} \circ g^{-1}$, where $\widehat{u}$ is the Poisson integral of $U \circ g^{-1}$. Then $u(z) = \widehat{u}(g(z)) = \widehat{u}(0)$. The boundaries $|z| = 1$ and the real axis correspond via

$$\mathrm{e}^{\mathrm{i}t} = g(\tau) = \frac{\tau - z}{\tau - \overline{z}} \quad (t \in [0, 2\pi], \tau \in \mathbb{R}).$$

Formally,

$$\mathrm{i}\mathrm{e}^{\mathrm{i}t}\mathrm{d}t = \frac{z - \overline{z}}{(\tau - z)^2}\mathrm{d}\tau$$

and so

$$\mathrm{d}t = \frac{2y}{|\tau - \overline{z}|^2}\mathrm{d}\tau.$$

Assuming the integrals do indeed transform in the way this suggests, we find that

$$u(x, y) = u(z) = \frac{y}{\pi}\int_{-\infty}^{\infty} \frac{U(\tau)}{(\tau - x)^2 + y^2}\,\mathrm{d}\tau.$$

This is, re-assuringly, the solution we obtained in 22.15 using the Fourier transform.

## Further examples of conformal mappings

We deliberately restricted our discussion of conformal mapping in Chapter 8 so as to focus on the guiding principles. With a view to potential applications we now extend our range of mappings.

**23.17 Mapping by trigonometric and hyperbolic functions.** Consider, as an example, the image of the semi-infinite strip $G = \{\, z : \operatorname{Im} z > 0,\ 0 < \operatorname{Re} z < \pi \,\}$ under

$$f \colon z \mapsto w = \cos z = \frac{1}{2}\left(\mathrm{e}^{\mathrm{i}z} + \mathrm{e}^{-\mathrm{i}z}\right).$$

It is far from obvious where $f$ sends $G$. However it is relatively easy to see what happens to the boundary lines: $f$ maps

| | | |
|---|---|---|
| the ray $\operatorname{Re} z = 0$, $\operatorname{Im} z \geqslant 0$ | to | the interval $[1, \infty)$, |
| the interval $[0, \pi]$ | to | the interval $[-1, 1]$, |
| the ray $\operatorname{Re} z = \pi$, $\operatorname{Im} z \geqslant 0$ | to | $(-\infty, -1]$. |

To obtain the last two of these, use the fact that $\cos(\pi + iy) = -\cosh y$. This shows us that the boundary of $G$ maps to the real axis. We conclude that $G$ itself maps either to $\Pi^+$ or to $\Pi^-$. The image of any point on the line $\mathrm{Re}\, z = i\pi/2$, $\mathrm{Im}\, z > 0$ has positive imaginary part. This confirms that the image of $G$ is the open upper half-plane.

In the same way, $\cosh z$ will map a semi-infinite strip $0 < \mathrm{Re}\, z < \pi$, $\mathrm{Im}\, z >$ to $\Pi^+$.

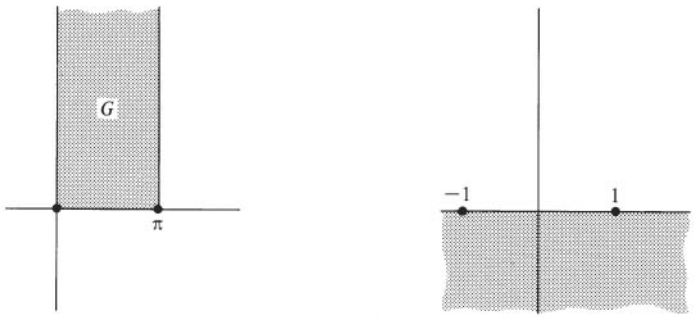



**Figure 23.4** From a semi-infinite strip to a half-plane

**23.18 The Joukowski transformation.** We consider the simplest form of the **Joukowski transformation**,

$$z \mapsto w = \tfrac{1}{2}(z + z^{-1}).$$

This satisfies $2wz = z^2 + 1$ and hence is easily seen to be given equivalently by

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

It is holomorphic in $\widetilde{\mathbb{C}}$ except at 0 and $\infty$, and conformal except at $\pm 1$, where angles are doubled.

Suppose that $w = u + iv$ is the image of $z = re^{i\theta}$, so that

$$u = \tfrac{1}{2}(r + r^{-1})\cos\theta, \quad v = \tfrac{1}{2}(r - r^{-1})\sin\theta.$$

The image of the half-line $\arg z = \mu$ is

$$\frac{u^2}{\cos^2\mu} - \frac{v^2}{\sin^2\mu} = 1,$$

which is a hyperbola. The image of the circle $|z| = \rho \neq 1$ is the ellipse

$$\frac{u^2}{\frac{1}{4}(\rho+\rho^{-1})^2} + \frac{v^2}{\frac{1}{4}(\rho-\rho^{-1})^2} = 1.$$

When $\rho = 1$, the image is the segment $[-1, 1]$ of the real axis.

The Joukowski transformation and variants of it have a distinctive feature which make them of special interest in fluid dynamics: they map certain circles to (models of) aerofoil shapes (see Fig. 23.5). This enables the lift on a model of an aircraft wing to be estimated.

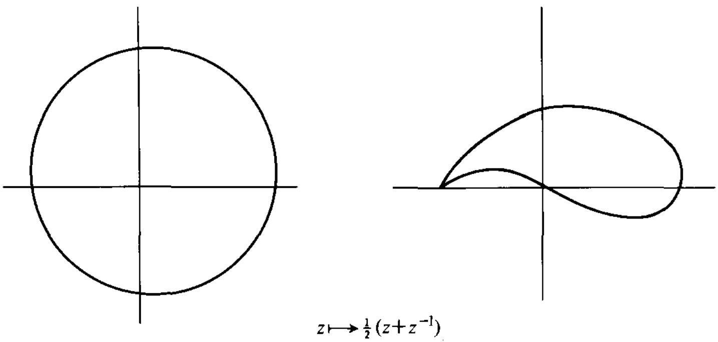



**Figure 23.5** Mapping an aerofoil

The Joukowski transformation provides another approach to the mapping of an semi-infinite strip, first considered in 23.17.

**23.19 Example (semi-infinite strip again).** To find a conformal map of $H = \{ z : \operatorname{Im} > 0,\ 0 < \operatorname{Re} z < \pi \}$ onto a half-plane.

**Stage 1** Let $w = e^{iz}$. Then (see 8.9) $|w| = e^{-\operatorname{Im} z}$ and $\arg w = \operatorname{Re} z \pmod{2\pi}$. Hence $H$ is mapped conformally onto

$$G = \{ w : 0 < |w| , 1,\ 0 < \arg w < \pi \}.$$

**Stage 2** We are now on familiar ground! Stages 1 and 2 of Example 8.14 supply the map $w \mapsto \zeta = ((w+1)/(w-1))^2$ from $G$ onto a half-plane. More directly, we can use the Joukowski transformation $w \mapsto \zeta = \frac{1}{2}(w + w^{-1})$. Think of $G$ as

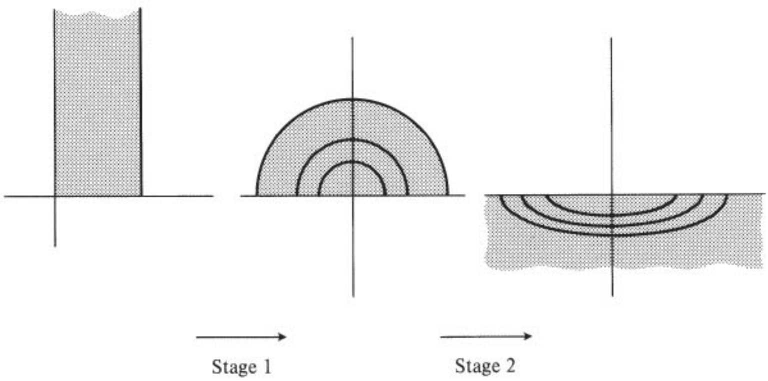



**Figure 23.6** From a semi-infinite strip to a half-plane

the union of semicircular arcs $w = re^{i\theta}$ $(0 < \theta < \pi)$, for $0 < r < 1$. Such an arc is mapped to the portion of the ellipse

$$\frac{u^2}{\frac{1}{4}(\rho + \rho^{-1})^2} + \frac{v^2}{\frac{1}{4}(\rho - \rho^{-1})^2} = 1$$

lying in the open lower half-plane. (Note that $\operatorname{Im}\zeta = \frac{1}{2}(r - r^{-1})\sin\theta < 0$.) As $r$ varies, the union of these image curves covers $\Pi^-$.

We conclude that $H$ is mapped conformally onto the open lower half-plane by the composite transformation

$$z \mapsto w = \tfrac{1}{2}(e^{iz} + e^{-iz}) = \cos z.$$

To reconcile this with 23.17, first rotate the given horizontal strip to the vertical strip $\{ z : \operatorname{Im} z > 0,\ -\pi < \operatorname{Re} z < 0 \}$ using the map $z \mapsto iz$ and remember that $\cosh z = \cos iz$.

**23.20 The Dirichlet problem for a semi-infinite strip.** Let

$$G = \{ z : 0 < \operatorname{Re} z < \pi,\ \operatorname{Im} z > 0 \}.$$

We seek a function $u$ such that

(i) $u$ is continuous on $\overline{G}$ except at 0 and is harmonic in $G$;
(ii) $u(z) = -1$ when $\operatorname{Re} z = \pi$ $(\operatorname{Im} z \geqslant 0)$ and when $\operatorname{Im} z = 0$ $(0 < \operatorname{Re} z \leqslant \pi)$;
(iii) $u(z) = 0$ when $\operatorname{Re} z = 0$ $(\operatorname{Im} z > 0)$.

As indicated in 23.17, $z \mapsto w = \cos z$ maps $\overline{G}$ one-to-one onto the closed lower half-plane, with $G$ mapped conformally onto the open lower half-plane. The boundary is mapped to the real axis, with $g(0) = 1$ and $g(\pi) = -1$. Define

$$\widehat{u}(w) = \frac{1}{w} \arg(w-1), \quad \text{where arg takes values between } -\pi \text{ and } \pi.$$

As the real part of a holomorphic branch of the logarithm, $\widehat{u}$ is harmonic. It takes the value 0 on $(1, \infty)$ and $-1$ on $(-\infty, 1)$. A suitable choice for $u$ is therefore

$$u(z) = \frac{1}{\pi} \arg(\cos z - 1) \quad (z \in \overline{G}).$$

In applications, maps of exterior regions are important. For example, in 2-dimensional fluid flow problems, we may want to analyse the flow past some obstacle. Our final examples treat problems of this sort.

**23.21 Example (semi-infinite slit).** Let

$$G = \mathbb{C} \setminus [0, \infty) = \{ z : 0 < \arg z < 2\pi \}.$$

To map $G$ to a half-plane, we need a square-root function. We cut the plane along the excluded slit and in the cut plane define

$$f : z \mapsto z^{1/2} = |z|^{1/2} e^{i\theta/2} \quad (0 < \theta < 2\pi).$$

Then $f$ maps $G$ conformally onto $\Pi^+$. See Fig. 23.7.

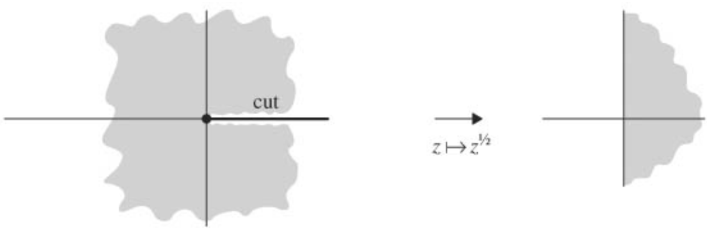



**Figure 23.7** semi-infinite slit

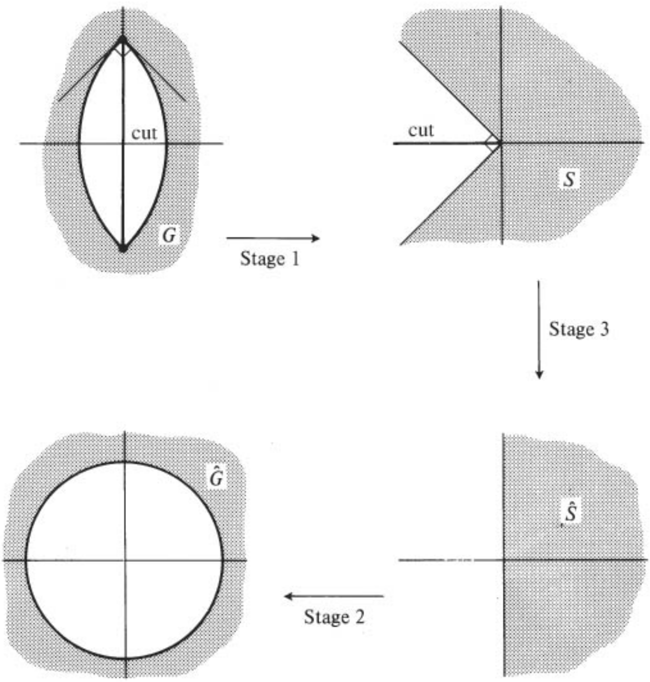



**Figure 23.8** Mapping the outside of a lozenge

**23.22 Example (region exterior to a lozenge).** To find a conformal mapping of the region $G$ exterior to both the circles $|z \pm 1| = \sqrt{2}$ onto the region $G$ exterior to the unit circle.

**Stage 1** The given region is bounded by circular arcs meeting orthogonally at $\pm\mathrm{i}$ (see Fig. 23.8). Take

$$g\colon z \mapsto w = \frac{z - \mathrm{i}}{z + \mathrm{i}};$$

$g$ is conformal except at $-\mathrm{i} \notin G$. The image of $G$ is a sector $S$ of angle $3\pi/2$ (see 8.3 and 8.10). The segment $[-\mathrm{i}, \mathrm{i}]$ bisects the angle between the arcs at $\mathrm{i}$ so its image bisects the complement of the image sector. Since $g(0) = -1$, the image is as shown in the figure.

**Stage 2** Working from the other end, we can realize our target region, namely $\{\, \tau : |\tau| > 1 \,\}$, as the image under the conformal map

$$h\colon \zeta \mapsto \tau = \frac{\zeta + 1}{\zeta - 1}$$

of the right half-plane $R = \{ \zeta : |\zeta - 1| < |\zeta + 1| \}$.

**Stage 3** To transform $S$ onto $R$, we seek to multiply angles at 0 by $2/3$. This suggests taking an appropriate holomorphic branch of $w \mapsto \zeta = w^{2/3}$.

We start in the $z$-plane with the plane cut along $[-\mathrm{i}, \mathrm{i}]$ (chosen to map to a suitable cut, along the negative real axis in the $w$-plane). In the cut $z$-plane, there exists a holomorphic branch $k$ of $[\![ ((z - \mathrm{i})/(z + \mathrm{i}))^{2/3} ]\!]$. The map we finally require is $f = h \circ k \circ g$. It is given by $f \colon z \mapsto \tau$, where

$$\left(\frac{z - \mathrm{i}}{z + \mathrm{i}}\right)^2 = \left(\frac{w + 1}{w - 1}\right)^3.$$

Although we have now mapped regions of many shapes there is one notable omission. We have not considered regions with polygonal boundaries.

**23.23 Mapping polygons.** Because of the way they act on arcs and circlines, Möbius transformations and exponentials are of most use for mapping regions whose boundaries are made up of curves of this sort. However their scope is, even so, limited. For example, to map regions with polygonal boundaries, it is necessary to introduce the **Schwarz–Christoffel transformation**

$$z \mapsto \int_0^z (\zeta - z_1)^{-k_1} (\zeta - z_2)^{-k_2} \ldots (\zeta - z_n)^{-k_n} \, \mathrm{d}\zeta.$$

Different paths from 0 to $z$ give different values to the integral, so we have a multifunction. Working with a suitable holomorphic branch, and suitably specified $k_1, \ldots, k_n$, we arrive at a map onto a disc of the polygon with vertices at $z_1, \ldots, z_n$. An introduction to this important but advanced material can be found, for example, in [**29**].

## Exercises

**Exercises from the text.** Fill in the details of the proof of the Transfer lemma (23.13).

23.1 Check that each of the following functions is harmonic on the indicated set, find a holomorphic function $f$ of which it is the real part, and also find a harmonic conjugate $v$:

| | |
|---|---|
| (i) $x^2 - y^2 - x$ | (on $\mathbb{C}$), |
| (ii) $x - y(x^2 + y^2)^{-1}$ | (on $\mathbb{C} \setminus [0, \infty)$), |
| (iii) $\sin(x^2 - y^2)\mathrm{e}^{-2xy}$ | (on $\mathbb{C}$), |
| (iv) $\log(x^2 + y^2)^{1/2}$ | (on the open first quadrant). |

23.2 Suppose that $u$ is harmonic in $\mathbb{C}$. Show that, for $0 \leqslant r < R$,

$$u(r\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{(R^2 - r^2)}{(R^2 - 2rR\cos(\theta - y) + r^2)} u(R\mathrm{e}^{\mathrm{i}t})\,\mathrm{d}t.$$

Assume in addition that $u \geqslant 0$. Deduce that

$$\frac{R-r}{R+r}u(0) \leqslant u(r\mathrm{e}^{\mathrm{i}\theta}) \leqslant \frac{R+r}{R-r}u(0).$$

Hence show that a bounded function which is harmonic in $\mathbb{C}$ is necessarily constant (cf. Liouville's theorem (13.3)).

23.3 Suppose that $f$ is holomorphic inside and on $\gamma(0;1)$. Let $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$.

(i) Evaluate the integrals

$$\frac{1}{2\pi}\int_0^{2\pi} (u(\mathrm{e}^{\mathrm{i}t}) \pm v(\mathrm{e}^{\mathrm{i}t}))\mathrm{e}^{-\mathrm{i}nt}\,\mathrm{d}t \quad (n = 0, 1, \ldots)$$

in terms of the coefficients of the Taylor series for $f$.

(ii) Deduce that, if $f(0)$ is real,

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \left(\frac{\mathrm{e}^{\mathrm{i}t} + z}{\mathrm{e}^{\mathrm{i}t} - z}\right) u(\mathrm{e}^{\mathrm{i}t})\,\mathrm{d}t \quad (|z| < 1).$$

23.4 Let $f$ be holomorphic in $A := \{ z : |z| > R \}$, with Laurent expansion $f(z) = \sum_{n=-\infty}^{\infty} c_n z^n$ in $A$. Let $\gamma$ be a positively oriented contour [closed path] in $A$ enclosing $0$. Prove that

$$\int_\gamma f(z)\,\mathrm{d}z = 2\pi \mathrm{i} c_{-1}.$$

(This variant on Lemma 18.2 is useful in connection with the complex potential associated with an obstacle whose boundary is described by $\gamma^*$.)

23.5 Consider Blasius's theorem as stated in 23.1. Calculate $F = (X, Y)$ and $M$ for the two-dimensional fluid flows given by the following complex potentials. (Hint: remember Lemma 18.1; expand the integrand and look only at the term in the integrand in $z^{-1}$.)

(i) $w(z)$ such that

$$\frac{\mathrm{d}w}{\mathrm{d}z} = U\left(z + \frac{a^2}{z}\right) - \frac{\mathrm{i}\Gamma}{2\pi z} \quad (U, a, \Gamma \text{ constants}),$$

where $\gamma = \gamma(0; r)$. (Here $w$ is the complex potential associated with flow past a circular cylinder of radius $a$ on which is superimposed a purely rotational flow round the cylinder with circulation $\Gamma$. The term $z + \frac{a^2}{z}$ arises from an application of the Milne–Thomson circle theorem mentioned in 23.2.

(ii) $w(z) = (z - b)^2 + \left(\frac{a^2}{z} - b\right)^2$, and $\gamma$ is such that $0$ lies inside $\gamma$ and $b$ lies outside.

(iii) $w(z) = U(z \cos\alpha - \mathrm{i}\sqrt{z^2 - a^2}\sin\alpha)$, where $U$ and $\alpha$ are constants and the plane is cut along $[-a, a]$. (Here $w(z)$ is the complex potential for the flow arising when a flat plate of length $2a$ is inserted into a uniform flow, $\alpha$ being the angle of inclination.)

23.6 Suppose that $W(z) = Uz + g(z)$ is the complex potential representing flow past a body whose cross-section is bounded by a contour $\gamma$, so that $\gamma^*$ is a streamline. Prove that the force $F$ on the body is zero. (Hint: recall Exercise 23.4.)

23.7 Find the image of the strip $\{0 < \operatorname{Re} z < \pi/2\}$ under the map $z \mapsto \operatorname{cosec}^2\left(\frac{\pi}{4} + \frac{z}{2}\right)$.

23.8 Show that the image under the transformation $g\colon z \mapsto z + a^2/z$ of the region exterior to the circle $|z| = a$ is $\mathbb{C} \setminus [-2a, 2a]$. Find, as the branch of a multifunction in a suitable cut plane, a conformal inverse for $g$.

23.9 Let $G$ be the region exterior to the pair of circular arcs through $\pm c$ and subtending angles of $\pm 2\pi/3$. Show that $G$ is mapped conformally to the region exterior to the unit circle by a map $f\colon z \mapsto \zeta$, where

$$\left(\frac{\zeta - 1}{\zeta + 1}\right)^4 = \left(\frac{z - c}{z + c}\right)^3;$$

you should specify carefully how the $z$-plane should be cut and how $f$ is obtained as a multifunction branch.

23.10 Find a conformal mapping of $G = \{\, z : 0 < \arg z < 3\pi/2 \,\}$ onto a strip. Hence find a function which is continuous on $\overline{G}$ except at 0, which is harmonic in $G$, and which is such that $u(x, 0) = 1$ $(x > 0)$ and $u(0, y) = 0$ $(y < 0)$.

# Appendix: new perspectives

Virtually everything so far in this book dates from well before the end of the 19th century. This brief appendix gives glimpses of a selection of more contemporary developments: the Prime number theorem, the Bieberbach conjecture, and Julia sets. Our intention is to convey just a little of the flavour of these topics, and the way in which theory presented in earlier chapters feeds into them. We hope you will be tempted to find out more by dipping into the specialized texts cited.

## The Prime number theorem

Andrew Wiles' proof of Fermat's last theorem in 1994 captured the imagination of many people, by no means all within the mathematical community. It drew attention to the fact that the statements of many famous problems in number theory are comprehensible, and intriguing, to non-mathematicians. Since Euclid proved that there are infinitely many primes, mathematicians have sought to understand how the prime numbers are distributed amongst the natural numbers. A way to do this is to obtain information, for real $x$, about the number, $p(x)$, of primes $\leqslant x$. A key result here is the Prime number theorem, which gives an asymptotic estimate of how fast $p(x)$ grows: $p(x)\log x/x \to 1$ as $x \to \infty$. In this section we give a skeleton of a proof of this theorem (dating from 1896) which makes crucial use of techniques from complex analysis. The details can be found in [**9**].

**A.1 The functions $\zeta(z)$ and $\Gamma(z)$.** The Riemann zeta function is defined for $\operatorname{Re} z > 1$ by

$$\zeta(z) := \sum_{n=1}^{\infty} \frac{1}{n^z}.$$

Here $n^{-z} := \mathrm{e}^{-z\log n}$ (as always, log denotes the logarithm to base e). Since $|n^{-z}| = n^{-\operatorname{Re} z}$, the series defining $\zeta(z)$ certainly converges when $\operatorname{Re} z > 1$; by Exercise 14.8 it is holomorphic there. It is easy to see that

$$(1-2^{-z})\zeta(z) = 1 + \frac{1}{3^z} + \frac{1}{5^z} + \frac{1}{7^z} \ldots,$$

$$(1-3^{-z})(1-2^{-z})\zeta(z) = 1 + \frac{1}{5^z} + \frac{1}{7^z} + \frac{1}{11^z} + \ldots,$$

and so on. This enables us to deduce that

$$\zeta(z) = \frac{1}{\lim_{N\to\infty} \prod_{k=1}^{N} (1 - p_k^{-z})},$$

where $p_1, p_2, \ldots$ is an enumeration of the primes ($p_1 < p_2 < \ldots$).

A key ingredient in the proof of the Prime number theorem is the method of analytic continuation discussed fleetingly in 15.12. The zeta function can be analytically continued in a variety of ways to a larger domain than the one, $S := \{ z : \operatorname{Re} z > 1 \}$, on which the defining series converges. Here we shall do this by relating it to another fundamental function. This is the gamma function:

$$\Gamma(z) := \int_0^\infty t^{z-1} \mathrm{e}^{-t} \, \mathrm{d}t \quad (\operatorname{Re} z > 0).$$

Proving holomorphy of a function such as this which is defined by an infinite integral is a technical business, beyond the scope of this book. However those with sufficient proficiency in integration theory can show quite easily that $\Gamma(z)$ is holomorphic for $\operatorname{Re} z > 0$. For $n = 0, 1, 2 \ldots$, $z \mapsto \Gamma(z + n + 1)$ is then holomorphic for $\operatorname{Re} z > -n - 1$. By integrating the defining integral by parts, we see that, for any $n \geqslant 1$,

$$\Gamma(z) = \frac{\Gamma(z + n + 1)}{z(z+1)\ldots(z+n-1)(z+n)} \quad (\operatorname{Re} z > 0).$$

The function on the right-hand side is holomorphic for $\operatorname{Re} z > -n$, except at non-positve integers. We deduce that it is possible to continue $\Gamma(z)$ analytically to a function (also denoted $\Gamma(z)$) holomorphic in $\mathbb{C}$, except at the non-positive integers, $0, -1, -2, \ldots$; at these points $\Gamma(z)$ has simple poles.

The gamma function has a number of interesting properties, of which we note the following:

- $\Gamma(n+1) = n!$ $(n = 0.1, \ldots)$;
- $\Gamma(z)\Gamma(1-z) = \pi \operatorname{cosec} \pi z$;
- $\Gamma(z) = \lim_{n\to\infty} \dfrac{n!\, n^z}{z(z+1)\ldots(z+n)}$.

We now need the following non-trivial fact: there exists a function $I(z)$ holomorphic in $\mathbb{C}$ such that

$$I(z) := (\mathrm{e}^{2\pi \mathrm{i} z} - 1)\zeta(z)\Gamma(z) \quad \text{for } \operatorname{Re} z > 1.$$

Here $I(z)(\mathrm{e}^{2\pi\mathrm{i}z} - 1)$ is given by the complex integral $\int_0^\infty t^{z-1}(\mathrm{e}^t - 1)^{-1} \, \mathrm{d}t$. To see how the displayed identity comes about, expand $(\mathrm{e}^t - 1)^{-1}$ as a series in powers of $\mathrm{e}^{-kt}$ and integrate term-by-term.

Now consider $I(z)/((e^{2\pi iz}-1)\Gamma(z))$. It can be shown that $\Gamma(z)$ is never zero and that $\zeta(z) \neq 0$ at $2, 3, \ldots$. Our results on the relationship between zeros and poles in 17.12 and 17.13 show that the singularities of $\Gamma(z)$ at $0, -1, -2, \ldots$ are cancelled by the zeros of $(e^{2\pi iz}-1)$. Therefore the function on the right-hand side of the above display has only removable singularities at the points of $\mathbb{Z} \setminus \{1\}$; remove these singularities to obtain an analytic continuation of $\zeta(z)$ to $\mathbb{C} \setminus \{1\}$ (again denoted $\zeta(z)$); at $z = 1$ there is a simple pole, of residue 1.

Before moving on to discuss the Prime number theorem we digress to mention the **Riemann hypothesis**, one of the most tantalizing and challenging of the longstanding open problems in mathematics. This concerns the zeros of the (extended) zeta function. In a paper published in 1859 Riemann conjectured that all the non-real zeros of $\zeta(z)$ lie on the line $\mathrm{Re} = 1/2$. As Riemann already demonstrated, there is a tight connection between the zeros of $\zeta(z)$ and the properties of the function $p(x)$, so that a resolution of the conjecture would have a major impact on number theory. It was proved long ago by G.H. Hardy that $\zeta(\frac{1}{2} + it) = 0$ for infinitely many real values of $t$. With the advent of powerful computers, evidence in favour of the hypothesis has mounted. For example, the value of $b$ is known for which are there precisely $10^9 + 1$ zeros having imaginary part in $(0, b)$; the order of magnitude of $b$ is $10^8$. The real part of every one of these zeros is equal to $1/2$. However, information of this sort cannot decide the matter. The general result remains elusive.

**A.2 The Prime number theorem.** Let $p(x)$ denote the number of primes $\leqslant x$. The theorem asserts that

$$\text{(P)} \qquad \lim_{x\to\infty} \frac{p(x)\log x}{x} = 1.$$

The first step in the proof of the theorem is the reduction to an equivalent problem. Let

$$q(x) := \sum_{n \leqslant x} L(n),$$

where

$$L(n) := \begin{cases} \log p & \text{if } n = p^m \text{ for some prime } p \text{ and some } m = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

. Some careful real analysis, relating $p(x)$ and $q(x)$, shows that (P) holds if it can be proved that $q(x)/x \to 1$ as $x \to \infty$.

The function $q$ is monotonic non-decreasing on $(1, \infty)$ with a jump of size

$L(n)$ at $n = 2, 3, \ldots$. Then

$$\sum_{n=1}^{\infty} \frac{L(n)}{n^z} = z \int_0^{\infty} q(u) \mathrm{e}^{-zu} \, \mathrm{d}u.$$

[The intermediate steps here involve Stieltjes integration:

$$\sum_{n=1}^{\infty} \frac{L(n)}{n^z} = \int_1^{\infty} t^{-z} \, \mathrm{d}q(t) = \int_0^{\infty} \mathrm{e}^{-zu} \, \mathrm{d}q(\mathrm{e}^u) = \int_0^{\infty} q(\mathrm{e}^u) \mathrm{e}^{-zu} \, \mathrm{d}u,$$

by the substitution $t = \mathrm{e}^u$ and then integration by parts.]

Exercise 14.8 sought a proof that $\zeta(z)$ defines a holomorphic function for $\operatorname{Re} z > 1$. By 14.7, its derivative is given by term-by-term differentiation:

$$\zeta'(z) = \sum_{n=1}^{\infty} \frac{\log n}{n^{-z}} \quad (\operatorname{Re} z > 1).$$

It can then be shown that

$$\frac{\zeta'(z)}{\zeta(z)} = -\sum_{n=1}^{\infty} \frac{L(n)}{n^z}.$$

[This relies on a technique from combinatorics known as Möbius inversion.] The limiting behaviour of $q(x)/x$ is therefore linked via $\zeta'(z)/\zeta(z)$ to the properties of (the analytic continuation of) $\zeta(z)$. We therefore have a function $G$ satisfying

$$G(z) := -\frac{1}{z} \left( \frac{\zeta'(z)}{\zeta(z)} + \frac{z}{z-1} \right).$$

This is certainly holomorphic for $\operatorname{Re} z > 1$ and coincides there with

$$\int_0^{\infty} q(\mathrm{e}^u) \mathrm{e}^{-zu} \, \mathrm{d}u - \frac{1}{z-1}.$$

But we can say more. Some delicate analysis shows that $\zeta(z)$ is never zero on the line $\operatorname{Re} z = 1$; the starting point for an argument by contradiction comes from the discussion of counting zeros in 15.13. It is a consequence that $G(z)$ is in fact holomorphic for $\operatorname{Re} z \geqslant 1$. This gives control on the limiting behaviour of $G(z)$ as $\operatorname{Re} z$ tends to 1. The proof is completed with the application of some quite sophisticated integration theory (though not beyond the level of [**6**]).

## The Bieberbach conjecture

The celebrated Bieberbach conjecture shares with the Prime number theorem the property of being appealingly simple to pose. It was solved affirmatively by L. de Branges in 1984, having been postulated nearly 70 years earlier. The book [**30**] presents the solution and some of the multitude of ancillary results spawned by the conjecture.

**A.3 Introducing the Bieberbach conjecture.** We consider univalent functions on the unit disc. A holomorphic function is called **univalent** if it is one-to-one. As we proved in Chapter 16, $f$ necessarily has a non-zero derivative, and so is conformal, and also has a well-defined and conformal inverse. The class of univalent functions has many interesting properties. In particular, we may ask what can be said about the coefficients in their Taylor expansions.

Assume that $f \in \mathrm{H}(\mathrm{D}(0;1))$ is univalent. It is no loss of generality to assume that $f(0) = 0$ and that $f'(0) = 1$ (for the latter, remember that $f'(0) \neq 0$, so that we can replace $f$ by $f/f'(0)$. The Taylor series for $f$ takes the form

$$f(z) = z + a_2 z^2 + a_3 z^3 + \dots .$$

The **Bieberbach conjecture** is the statement that

$$|a_n| \leqslant n \quad (n = 2, 3, \dots).$$

Certainly the conjecture is true, with equality in the bound, for any **Koebe function**

$$K_\lambda(z) := \frac{z}{(1-\lambda z)^2} = \sum_{n=1}^{\infty} n \varepsilon^{n-1} z^n \quad \text{where } \lambda \text{ is a constant, } |\lambda| = 1.$$

These functions are, in fact the *only* functions for which $|a_n| = n$ for some (and hence for all) $n$.

The conjecture came to be postulated because of a corollary of a result due to Bieberbach. This result, which is of independent interest, is the following one.

**A.4 Bieberbach's area theorem.** Let $G$ be holomorphic for $|z| > 1$ and have Laurent expansion

$$G(z) = z + c_0 + \frac{c_1}{z} + \frac{c_2}{z^2} + \dots \quad (|z| > 1).$$

Then

$$\sum_{n=1}^{\infty} n\,|c_n|^2 \leqslant 1.$$

This inequality is proved by calculating the area enclosed by $\gamma(0;r)$ for $r > 1$. This area, which is certainly non-negative, can be shown to be

$$A(r) = \pi \left( r^2 - \sum_{n=1}^{\infty} n \left| c_{2n} \right| r^{2n} \right).$$

The conclusion follows if we let $r \to 1$.

**A.5 Bieberbach conjecture: the case $n = 2$ and beyond.** Let $f$ be as in A.3. It is not difficult to show that

$$F(z) := z + \frac{a_2}{2} z^2 + \dots$$

is holomorphic in $\mathrm{D}(0;1)$, univalent, and non-zero for $z \neq 0$. Finally, we can show that the area theorem applies to

$$G(z) := 1/F(1/z) = z - \frac{a_2}{2} z^{-1} + \dots$$

to give $|-a_2/2| \leqslant 1$.

A quite different technique, involving a differential equation, was introduced by K. Löwner in 1923 to treat the case $n = 3$. Variational methods were brought to bear on the problem in the 1930s and this led to the conjecture being established for $n = 4$ in 1955 and for $n = 6$ in 1972. Finally, twelve years later, de Branges proved the general case.

**A.6 The Koebe $\frac{1}{4}$-theorem.** Another famous result, stemming from the case $n = 2$ in the Bieberbach conjecture, is worth mentioning. This states that a univalent function $f$ on $\mathrm{D}(0;1)$ with $f(0) = 0$ and $f'(0) = 1$ is such that $\mathrm{D}(0; \frac{1}{4}) \subseteq f(\mathrm{D}(0;1))$. with Equality occurs for the same class of functions as gives equality in the Bieberbach conjecture.

## Julia sets and the Mandelbrot set

This final section concerns the geometric and topological features of the complex plane associated with dynamical systems whose evolution is governed by a sequence $p, f(p), f(f(p)), \dots$ in $\mathbb{C}$. Such systems describe a variety of phenomena in many areas of physics and the life sciences, in particular those exhibiting chaotic behaviour.

**A.7 Iterating complex functions.** Let $f$ be a complex-valued function and fix $p \in \mathbb{C}$. We may then form the sequence of **iterates**

$$p, f(p), f(f(p)), f(f(f(p))), \ldots,$$

so obtaining a complex sequence $\{z_n\}$, where $z_0 = p$ and $z_{n+1} = f(z_n)$. We also write $z_n$ as $f^{[n]}(p)$; of course, this notation must not be confused with that for higher-order derivatives.

As a very simple case, let $f(z) = z^2$. Then $z_n = p^{2n}$. There are three possible outcomes as $n \to \infty$: if $|p| < 1$, then $z_n \to 0$; if $|p| > 1$, then $|z_n| \to \infty$; if $|p| = 1$ the point $z_n$ moves forever on the unit circle, converging only if $p = 1$. The unit circle divides the plane into two regions separated by the unit circle: a starting value $p$ in one results in $z_n$ being 'attracted' to 0, and a starting value in the other results in 'repulsion'.

Let $f$ be a function holomorphic in $\mathbb{C}$. If the points

$$\alpha, \ldots, f(\alpha), \ldots, f^{[q-1]}(\alpha)$$

are distinct, but $f^{[q]}(\alpha) = \alpha$, then we say the points form a $q$**-cycle** and that each point $f^{[k]}(\alpha)$ $(k = 0, \ldots, q-1)$ is a **periodic point**. If $q = 1$, we have a **fixed point**. Once $z_n = f^{[n]}(p)$ reaches a periodic point, then thereafter it cycles indefinitely through the points of the cycle. An easy exercise using the Chain rule shows that the derivative $f^{[q]\prime}$ takes the same value at each point of a $q$-cycle. $\alpha$ is **attracting** (or **an attractor**) if $|f^{[q]\prime}(\alpha)| < 1$ and **repelling** if $|f^{[q]\prime}(\alpha)| < 1$.

**A.8 Julia sets.** We shall henceforth restrict attention to iteration by quadratic functions $F_c := z^2 + c$. (More generally, we might consider $f(z) = az^2 + bz + c$ $(a \neq 0)$. However a transformation $\phi\colon z \mapsto az + \frac{b}{2}$ yields a function $\phi \circ f \circ \phi^{-1}$ of the form $F_c$, and the conjugate functions $f$ and $F_c$ exhibit the same behaviours.) How $\{z_n\}$ behaves turns out to be critically dependent on the choice of $c$. Let $r_c$ be the non-negative root of the equation $x^2 + c = x$. We then define, for any $c$,

$$E_c := \{\, p : |F_c{}^{[n]}(p)| \to \infty \,\} \qquad \text{(the \textbf{escape set})},$$
$$K_c := \mathbb{C} \setminus E_c \qquad \text{(the \textbf{keep set})}.$$

The following facts can be established:

- $z \in E_c$ $(K_c)$ if and only if $-z \in E_c$ $(K_c)$;
- $z \in E_c$ $(K_c)$ implies $f(z) \in E_c$ $(K_c)$;
- $K_c \subseteq \overline{\mathrm{D}}(0; r_c)$;

- $E_c$ is open and connected;
- $K_c$ is closed and is simply connected (informally, it has no 'holes' in it)
- $K_c$ is connected if and only if $0 \in K_c$ (the **Fatou–Julia theorem**).

For a periodic point $\alpha$ of $F_c$ we have:

- if $\alpha$ is an attractor then it is an interior point of $K_c$ (that is, there exists $r > 0$ such that $D(\alpha; r) \subseteq K_c$);
- if $\alpha$ is repelling then $\alpha$ is in the boundary $\partial K_c$ of $K_c$.

Finally, we define the **Julia set** of $F_c$ to be the boundary $\partial K_c$ of the associated set $K_c$ and the filled-in Julia set to be $K_c \cup \partial K_c$. A substantial theorem of complex analysis allows the Julia set to be characterized as precisely the set of points on which $\{f^{[n]}\}_{n \geqslant 0}$ does not form a normal family of functions. Loosely, a family $\mathcal{F}$ of holomorphic functions is normal at a point $z$ if on some $D(z; r)$ any sequence in $\mathcal{F}$ either has a subsequence converging uniformly on compact subsets to a holomorphic limit or a subseqeuence which, uniformly on compact subsets, tends to infinity. A discussion of normal families can be found in [**8**].

We may view the Julia set for a given value of $c$ as a 'curve' dividing the plane into two. Fig. A.1 shows one example. As Mandelbrot discovered, Julia sets are generally highly complex. They exhibit the property characteristic of fractals: by zooming in on a portion of the curve, however minute, and iterating by $F_c$ on this portion repeatedly, the entire curve is generated. At one extreme the Julia set can be a perfect circle, or a fractally deformed circle; at the other, it may fragment into a multitude of tiny flecks (called Fatou dusts), with $K_c$ having no interior points at all. Part of the fascination of the subject of fractal curves such as these is their enormous variety and their beauty. See [**32**] for a wealth of computer-generated illustrations in colour.

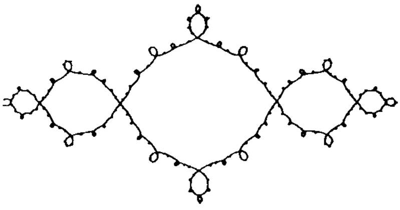


**Figure A.1** The Julia set for $z^2 - 1$

**A.9 The Mandelbrot set.** The **Mandelbrot set** $M$ was introduced in Exercise 3.16 and some elementary topological properties of it were presented. Here we define it by

$$M := \{ c \in \mathbb{C} : F_c^{[n]}(0) \text{ does not tend to infinity} \}.$$

That is, $c \in M$ if and only if $0 \in K_c$, the keep set for the function $F_c\colon z \mapsto z^2 + c$. The Fatou–Julia theorem characterizes $M$ alternatively in terms of the topological structure of the keep sets:

$$M = \{ c \in \mathbb{C} : K_c \text{ is connected} \}.$$

The set $M$ is a closed subset of $\overline{\mathrm{D}}(0;2)$, and so is compact. It is shown in Fig. A.2, for values of $c$ in the range $-2 \leqslant \operatorname{Re} c \leqslant 1$, $-1.5 \leqslant \operatorname{Im} c \leqslant 1.5$. As the figure suggests, $M$ is symmetric about the real axis, which it intersects in the interval $[-2, 1/4]$. It is also true, but far from obvious, that $M$ is connected and without holes. Large-scale pictures of $M$ reveal very clearly both the fractal nature of its boundary and presence of infinitely many hair-like branching filaments. The connectedness of $M$ relies on the existence of these filaments.

**Figure A.2** The Mandelbrot set

The significance of the Mandelbrot set lies in the way that the evolution of the system described by $F_c$, and the associated Julia set, is critically dependent on where $c$ lies relative to $M$. For example, for $c \in M$, the filled-in Julia set consists of a warped and distorted disc with a fractal boundary, and a great variety of shapes are possible. Moreover, the possible behaviours of the system described by $F_c$ can be classified according to where in $M$ the point $c$ lies. For example:

- for $c$ in the cardioid body of $M$, there is a single finite attractor;
- for $c$ in a 'bud', there is a 3-point attracting cycle;
- for $c$ lying on an antenna of $M$, the Julia set degenerates into a scatter of flecks and there is no attractor in $\mathbb{C}$.

Out beyond $M$ the filled-in Julia set becomes a set of 'islands', and its complement fails to be simply connected.

We saw in Chapter 23 that conformal mapping may be used to make problems in fluid dynamics more tractable. Likewise, conformal mapping can transform complicated complex dynamical systems to more tractable ones.

There is a deep result of A. Douady and J.H. Hubbard showing that, for any $c$, there is an invertible conformal mapping $\phi_c$ of some $\mathrm{D}(\infty; r)$ (in $\widetilde{\mathbb{C}}$) such that

$$\phi_c \circ F_c \circ \phi_c^{-1} = F_0,$$

and $\phi_c$ maps $E_c$ onto the complement of $\overline{\mathrm{D}}(0;1)$, for any $c \in M$. Moreover, for $c \notin M$, the map $c \mapsto \phi_c(c)$ is an invertible conformal mapping of the complement of $M$ onto the complement of $\overline{\mathrm{D}}(0;1)$. Observe that we are dealing here with regions which are not simply connected; compare with the Riemann mapping theorem (see 16.10). (This last result can be used to establish that $M$ is connected.) The results above reinforce the impression that, however chaotic the system described by $F_c$ may be near the origin, it is relatively easy to analyse out towards infinity.

The effect of a change of variable $z \mapsto \zeta = 1/z$ is to transform $F_c$ to $R_c$, given by $R_c(\zeta) = \zeta^2/(1 + c\zeta^2)$. Then $R_c$ can, locally, be transformed to $R_0$ by a map $\Phi_c$ analgous to $\phi_c$ above. Then $\Phi_c$ has an expansion

$$\Phi_c(\zeta) = \zeta + a_2(c)\zeta^2 + a_3(c)\zeta^3 + \dots \quad (z \in E_0).$$

An intriguing characterization of $M$ was obtained by F. v. Haeseler:

$$M = \{\, c \in \mathbb{C} : |a_n(c)| \leqslant n \ (n = 2, 3, \dots)\,\}.$$

Compare this with the Bieberbach conjecture!

# Bibliography

## For mathematical background material

1. T. M. Apostol, *Mathematical analysis* (1st edn), Addison–Wesley (1957) (for topology in Euclidean space and also for a Riemann integral approach to transform theory).
2. T. M. Apostol, *Mathematical analysis* (2nd edn), Addison–Wesley (1974).
3. R. G. Bartle and D. R. Sherbert, *Introduction to real analysis* (3rd edn), Wiley (2000).
4. K. G. Binmore, *Mathematical analysis: a straightforward introduction* (2nd edn), Cambridge University Press (1982).
5. M. Hart, *A guide to analysis*
6. H. A. Priestley, *Introduction to integration*, Oxford University Press (1997).
7. W. A. Sutherland, *Introduction to metric and topological spaces*, Oxford University Press (1975).

## Supplementary and collateral reading

The books [**20**] and [**21**] are especially recommended, for their treatment of the historical background to complex function theory and in particular Cauchy's contributions. Several of the books cited adopt a more geometric and topology approach to complex analysis than we have done and in particular do justice to some important topics, such as simple connectedness and Riemann surfaces, that we have treated very sketchily.

8. L. V. Ahlfors, *Complex analysis* (3rd edn), McGraw–Hill (1978).
9. R. B. Ash, *Complex variables*, Academic Press (1971).
10. A. F. Beardon, *Complex analysis*, Wiley (1979) (subtitled *the argument principle in analysis and topology*).
11. A. F. Beardon, *A primer on Riemann surfaces*, Cambridge University Press (1984).
12. R. B. Burckel, *An introduction to classical complex analysis*, Vol. I Birkhäuser Verlag (1979).
13. L. Hahn, *Complex numbers and geometry*, Mathematical Association of America (1994).
14. E. Hille, *Analytic function theory* (2nd edn) Chelsea (1973) (for triangulation of a polygon).
15. K. Hoffman, *Banach spaces of analytic functions*, Prentice–Hall (1962).

16. F. C. Kirwan, *Complex algebraic curves*, Cambridge University Press (1992).
17. C. Kosniowski, *A first course in algebraic topology*, Cambridge University Press (1980) (for the Jordan curve theorem).
18. T. Needham, *Visual complex analysis*, Oxford University Press (1997) (a highly geometric approach, with links to physics).
19. W. Rudin, *Real and complex analysis* (3rd edn), McGraw–Hill (1987).
20. R. Remmert, *Theory of complex functions* (translated by R.B. Burckel), Springer (1991).
21. F. Smithies, *Cauchy and the creation of complex function theory*, Cambridge University Press (1997).
22. I. Stewart and D. Tall, *Complex analysis*, Cambridge University Press (1983).
23. E. C. Titchmarsh, *The theory of functions* (2nd edn), Oxford University Press (1939) (a renowned classic).

## For connections with applications

24. M. J. Ablowitz, *Complex variables: introduction and applications* (2nd edn), Cambridge University Press (1997).
25. D. J. Acheson, *Elementary fluid dynamics*, Oxford University Press (1990).
26. G. R. Grimmett and D. R. Stirzaker, *Probability and random processes* (3rd edn), Oxford University Press (2001).
27. T. W. Körner, *Fourier analysis*, Cambridge University Press (1988).
28. E. J. Watson, *Laplace transforms and applications*, Van Nostrand Reinhold (1981).
29. R. A. Silverman, *Complex analysis with applications*, Prentice–Hall (1974).

## For the Appendix: new perspectives

30. S. Gong, *The Bieberbach conjecture*, Studies in advanced mathematics **12**, American Mathematical Society and International Press (1999).
31. B. B. Mandelbrot, *The fractal geometry of nature*, W. H. Freeman (1982).
32. H. O. Peitgen and P. H. Richter, *The beauty of fractals*, Springer-Verlag (1986).
For a proof of the Prime number theorem, see [**9**].

# Notation index

*This page intentionally left blank*

# Index

Page numbers given in boldface refer to definitions and those in italic to exercises, with the latter overriding the former where definitions are given in exercises.